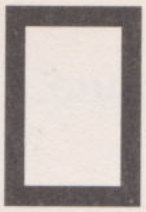
Prime Numbers

#3



NATIONAL GEOGRAPHIC

Prime Numbers

A long road to infinity

Enrique Gracián

Our mathematical world

Text by Enrique Gracián
English adaptation by Windmill Books and Vespa Design

ISBN 978-84-473-8766-3
Legal deposit: B 22737-2016

Printed in Spain by RODESA, Villatuerta-NAVARRA

Contents

Preface

When it comes to arithmetic, most numbers show what we might call 'good behaviour'. Even numbers always alternate with odd numbers, every third number is always a multiple of 3 and perfect squares obey an easily determinable law. So we can draw up a long list of numbers that do what they are supposed to do, no matter how large they are or where they fall. On the other hand, prime numbers are an unruly bunch. They appear wherever they want to, without prior warning, in an apparently chaotic fashion and without following any rule. And the worst of it is that they cannot be ignored – they are essential to arithmetic and to all mathematics.

Prime numbers are not a complicated phenomenon requiring years of mathematical study; indeed, they are taught in the early years of schooling. To understand what a prime number is, all you need to know is a counting system and the four basic arithmetical operations. However, they have been and continue to be one of the most astonishing challenges in the history of science. Anyone wishing to devote themselves to mathematics and cannot handle primes is lost, as they are ever present, always lying in wait, ever ready to pop up when least expected. And when primes appear, they do so in an inescapable and relentless fashion, marking out their territory and imposing their will.

Their influence is not only felt in the world of mathematics. Although we are unaware of it, prime numbers play a crucial role in our everyday lives, in the protection of our personal computers, in bank transactions or in the privacy of our mobile phone conversations. They are the cornerstones of computer security.

In a metaphorical sense, prime numbers are like a malicious virus, which, when it takes hold of a mathematician's mind, is very hard to eradicate. Euclid, Fermat, Euler, Gauss, Riemann, Ramanujan and a long list of the most famous mathematicians in history have fallen victim. Some were more or less successful in throwing it off, but all succumbed to the obsession with finding the 'magic formula', a rule that determines which prime will follow a particular number. However, no one has yet succeeded in finding it.

Prime numbers have left a trail of conjecture throughout the history of mathematics. In a way, it could be said that the history of prime numbers is a record of failure, but wonderful failures that, over time, have generated new theories, new paradigms and new milestones. In terms of mathematical creativity, prime numbers

have been a source of such richness that, although it is paradoxical to say so, it is a good thing they have not yet been mastered. And everything suggests that this situation will continue for a long time to come.

In preparing this book, we have tried to keep the level of explanation 'high', which means that the amount of mathematical knowledge needed to read it is 'low'. The inverted commas around both these adjectives indicate that these are relative concepts, and all the more so in the topic occupying us here. In any event, this book can be tackled by any reader who knows what numbers are and the basic operations that can be applied to them; its aim is to give the reader a concise guide to the world of prime numbers.

On the other hand and for those readers who have a more advanced knowledge of mathematics, we have tried to include information on particular historical processes that are essential for understanding the intricacies explored by the great mathematicians in their research into the problems posed by prime numbers.

In conclusion, as is made clear from the first chapter onwards, the concept of a prime number and the challenges presented by such numbers are simple and easy to explain, but the solutions proposed belong, in the majority, to the higher echelons of professional mathematics.

Chapter 1

At the Dawn of Arithmetic

Like everything else, prime numbers have an origin- a beginning to be found at the very start of counting systems. They first appeared along with the natural numbers, but very quickly stood out as a set of rather special numbers.

Nothing more natural than a natural number

"God made the first ten numbers; the rest is the work of man." This statement is attributed to Leopold Kronecker (1823–1891), a German mathematician. He was referring to the natural numbers that we use for counting, 1, 2, 3, 4, 5, etc. Kronecker was making the point that much of the mighty edifice of mathematics is built on the most basic, elementary arithmetic. But to assert that God gave us the first ten numbers is tantamount to saying – putting religion aside – that these numbers have always been present as part of Nature.

It would not be too much of a stretch to suppose that the process of counting started when humans gave up their hunter-gatherer way of life to start down the long road of arable farming and livestock breeding. With this change, many items, like grains of wheat, flocks of sheep or dairy herds, ceased to be something with a single immediate use and became commodities to be tallied, recorded and traded. This created the need for specific ways of counting. Let us imagine a shepherd who takes a flock out to graze. He needs to be sure that, when he returns, the same number enter the fold as left it. The most natural way of achieving this, if he does not possess a counting system, is to gather together a heap of pebbles and put one stone into a bag for each sheep going out. Then, when he returns, all he has to do is to take out a stone for each sheep going in and so check that the accounts tally. This is a primitive system of calculation – the word *calculation* comes from the Latin *calculus* meaning 'pebble'. The pebble system does not require the concept of a number. In modern mathematical parlance we would say that the shepherd establishes a bijective or biunivocal (one-to-one) relationship

NUMERICAL PERCEPTION

When the Chinese spoke of the 10,000 stars that there are in the sky, they were not claiming to have counted them all. It was simply a way of expressing a very large number. Someone else may think that a billion is a better number for expressing a numerically large concept. We need to bear in mind right from the start that our direct perception of a number is limited to five units. When someone shows all the fingers of one hand and three of the other, we quickly see a total of eight fingers, but this is almost a code. If we line up eight objects on a table we have to count them or visualise them as groups of smaller known quantities to say how many there are. Hence it is very difficult for us to picture a million units if we don't have an immediate reference. We know what it means to win a million pounds in the lottery because we know the value of money, and we quickly make a mental calculation of what we could buy with it. But there is big difference between this and having a clear idea of what it means to line up a million one pound coins (they would cover a distance of 22.5 km).

At a single glance our brains are able to recognise a maximum of five objects. For larger quantities, other strategies for counting them have to be used.

between the flock of sheep and the set of stones. We should note, however, that the mathematical concept of a biunivocal relationship between two sets was not established until the 19th century, and so it may seem paradoxical to consider this process of counting as one of the most natural. Therefore, when we state that something is 'natural' we are obliged, at least in this context, to make some clarifications.

We might imagine that a mental process that arises immediately without the need for previous thought is a natural one. But we cannot be sure that the counting system using a bag of stones does not require any previous reflection at all. In any event, it could equally be characterised by how easy it is to perform and whether the process is effective at achieving its purpose. Using the amount of thought re-

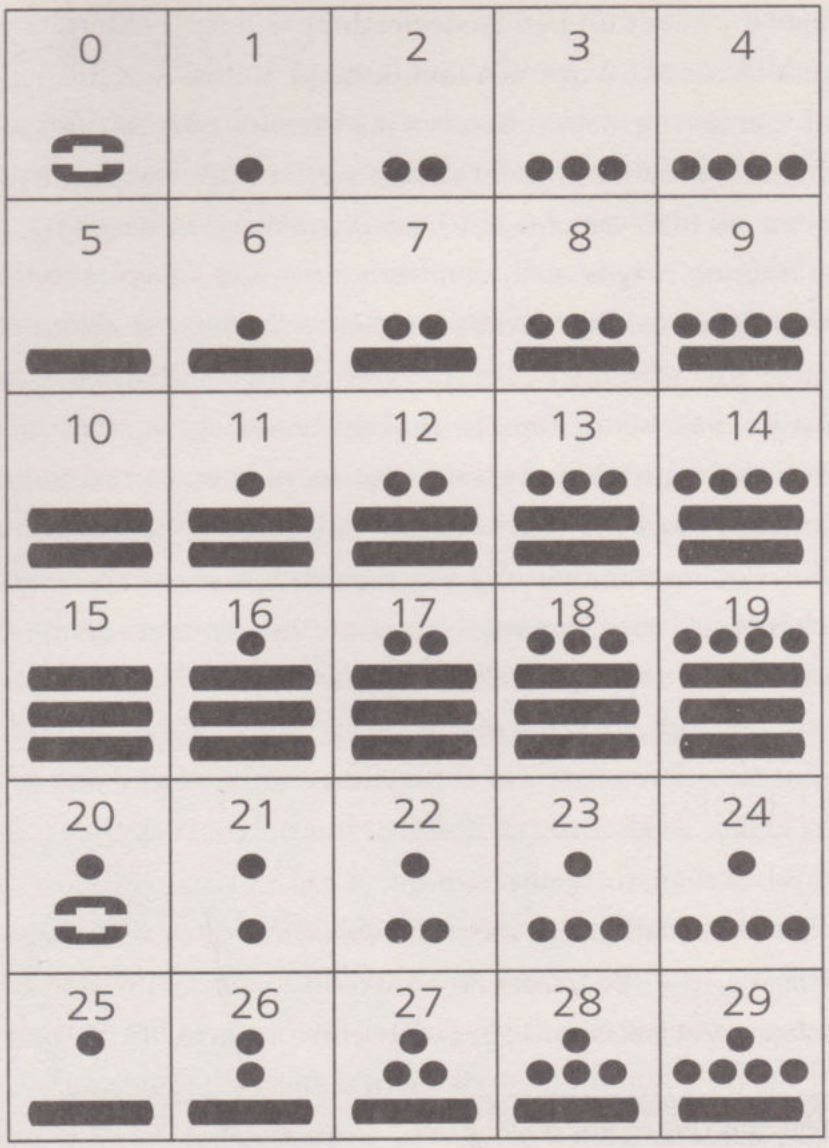
quired for a mental process to qualify its naturalness is not very effective. In this context we would do better to speak about levels of abstraction.

Introducing a counting system involves a powerful process of abstraction, so much so that many specialists consider that, together with learning language, it is one of the greatest mental leaps made by a human being in his or her whole life. When we say 'three' we may as well be referring to three sheep as to three stones, three houses, three trees or three of anything. If we had to use different words to quantify each object to which we refer, the embryonic agricultural society would have imploded right at the start. 'Three' is an abstract concept, a purely mental image that, in order to function in a social group, requires only one word and one sign as a means of communication.

Let us recall in passing that everyday language also involves processes of abstraction. When a child first learns the word 'chair', it usually refers exclusively to the object used for sitting on, but gradually he or she realises that the same word can refer not only to one high chair but also to many other objects around the house with the same function. The process of abstraction continues, and one day the word 'seat' crops up, a higher level of abstraction that not only incorporates all chairs but also benches, stools and anything that can be sat on.

Many people dislike mathematics, a dislike that they justify by declaring that the subject is too abstract, as if the process of abstraction were something artificial and unnatural. But that is not the case. If we did not use our capacity for abstraction we could not even devise a common language. Occasionally, abstract thought is also called impractical, but that is not true either. The more practical we want a method to be, the more elaborately and abstractly it must be designed. A good example of this is the positional numbering system we use every day – in a most 'natural' fashion. In a non-positional numbering system, a symbol representing a number has the same value whatever the position it occupies. For example, in the system of Roman numerals, the number 5, represented by the letter V, has the same value in the expressions XV, XVI or VII. However, if the Roman system had a positional numbering system like ours, the V would be equivalent to 5 units in the first instance, 50 in the second and 500 in the third.

Creating a positional numbering system was not exactly a simple task. It took more than 1,000 years to achieve. The history of numbers is long and exciting, but this is not our focus of attention here. Therefore, we shall consider for our purposes that the numbers are already there and that, in addition, we are familiar with the basic operations of addition, subtraction, multiplication and division.



The Mayan culture was one of the few civilisations of the ancient world to develop a positional numbering system. The Maya used just three symbols: a shell representing zero, a dot for each unit and a dash to express five units.

What is a prime number?

Let's take any number, for example, 12. We know that we can express this number in different ways as a product (·) of other numbers:

$$12 = 2 \cdot 6;$$
$$12 = 3 \cdot 4;$$
$$12 = 2 \cdot 2 \cdot 3.$$

We shall refer to these numbers henceforth as 'factors' and 'divisors'. Thus we can say that 3 is a factor of 12 or 3 is a divisor of 12. A divisor is small number that divides into a larger one, hence 3 divides into 12. Similarly we can say that 5 is a divisor of 20 because 5 divides into 20. What we mean by 'divide' in this context is that if we divide 20 by 5 we obtain a natural number, in this case 4, and the remainder of the division is zero.

The word 'factor' also has a precise meaning. It comes from the Latin *facere* meaning 'to make' or 'to produce'. In the expression $12 = 3 \cdot 4$, the number 3 is a factor because it is a number that 'produces' the number 12.

Accordingly, when asked, 'What are the divisors of 12?', we can reply that 2, 3, 4 and 6 are divisors of 12, because 12 divided by any of them yields a whole number. The set of divisors of any number will also include 1, as every number is divisible by one and also by itself. For example, if we ask what numbers divide into 18, we reply that 18 can be divided by 1, 2, 3, 6, 9 and 18.

SIGNS OF THE DEVIL

In the Dark Ages of European history, numerals or ciphers were considered to be the mysterious signs of a 'secret writing', and this is why even today coded messages can be called 'cipher messages'. Properly speaking, however, only messages in which the letters have been replaced by numbers should be called ciphers. When the first Arabic numerals were introduced to Europe in the columns of abacuses, abacus 'purists' substituted them with Roman numerals, refusing to allow the presence of these 'devilish symbols with which Satan had led the Arabs astray'. Even six centuries after the death of Pope Sylvester II in 1003, the Church ordered that his tomb be opened to see whether it still contained the demons that had inspired his interest in the Saracen science of numbers.

Gerbert of Aurillac was Pope Sylvester II, a rather mathematical pope.

Let's suppose that we ask ourselves the same question, but with the number 7. Looking for possible divisors we find that the only numbers that divide into 7 are 1 and 7 itself. Something similar happens with numbers like 2, 3, 5, 11 and 13. These numbers are called 'primes numbers'.

We are now in a position to give a precise definition of what a prime number is: a number is said to be prime when it is divisible only by 1 and itself.

These thoughts about natural numbers have involved the operations of multiplication and division. We have reached the conclusion that some numbers are special and, by finding a definition that covers them all, we have undertaken a process of abstraction. Having been given a name and a property that defines the numbers, they can now be studied in depth.

The fundamental theorem of arithmetic

Prime numbers are often referred to as the 'bricks' of mathematics, the 'atoms' of mathematics or the 'genetic code' of numbers. Houses are built with bricks, all the natural elements consist of atoms, and living things are defined by a genetic code. All these analogies are based on a common notion – primordial elements from which the whole of a system is produced. Let's take a look at the role played by prime numbers in mathematics.

We have seen that a number can be broken down into divisors or factors. Thus, the number 12 can be represented by $3 \cdot 4$. Remember that, when we speak of factors, we mean that we can produce 12 from the numbers 3 and 4. We know that we can also produce it from other numbers, for example:

$$12 = 2 \cdot 6 = 3 \cdot 4 = 2 \cdot 2 \cdot 3.$$

All these are factors of the number 12, and the process is known as factorisation. We remember that this was the criterion that allowed us to come up with a precise definition of what a prime number is: a number the only factors of which are 1 and itself. Accordingly, the only factors of a prime number such as 13 are 1 and 13:

$$13 = 1 \cdot 13$$

When one of the factors in a product is repeated, we place a superscript over the number to indicate how many times it is repeated. For example:

$$2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5;$$
$$3 \cdot 3 \cdot 3 \cdot 3 = 3^4.$$

This is what mathematicians call a 'power' and it is read as 2^5, two to the power five, and 3^4, three to the power four.

In a previous example we broke down the number 12 into three products with different factors: 2 and 6; 3 and 4; 2, 2 and 3. Only the last of these expressions consists solely of prime numbers.

Let's take a look at another example,:

$$20 = 2 \cdot 10 = 2 \cdot 2 \cdot 5 = 4 \cdot 5.$$

Only the product $20 = 2 \cdot 2 \cdot 5 = 2^2 \cdot 5$ contains prime factors only.

The question we now ask ourselves is, given a number at random, is it always possible to reduce it to prime factors? That is, can it be expressed as a product of numbers that are all primes? The answer is yes. Not only that, but it can only be done in one way. When we write the number 20 as a product of prime factors, $20 = 2^2 \cdot 5$, we do so in the only way possible – note that the order of the factors is not significant, as $2 \cdot 5 \cdot 2$ is the same as $5 \cdot 2 \cdot 2$. This is the theorem attributed to Euclid known as the 'fundamental theorem of arithmetic'. It states that "Any natural number can be reduced in a single way to a product of prime factors".

HOW TO FIND PRIME FACTORS

120	2
60	2
30	2
15	3
5	5
1	

To reduce a number to its prime factors, the best method involves placing the number concerned to the left of a vertical line. We then check whether the number is divisible by 2, 3, 5, etc., that is, by prime numbers, starting with the smallest. If it is divisible, we place the result of the division on the left and start again with this new number. The process is continued until the number on the left is one. The column on the right will then contain the prime numbers that factorise the number concerned.

Thus, when we write $24 = 2^3 \cdot 3$, we are saying that this is the only way to do so with prime factors. Hence, the name 'fundamental theorem' is totally justified as it is one of the foundations on which the whole of arithmetic rests. Furthermore, from this perspective prime numbers take on a vital dimension. Returning to the aforementioned comparisons, it could be said that $2^3 \cdot 3$ is the DNA of the number 24; that is a sequence consisting of the genes 2^3 and 3, or that 2 and 3 are the atoms making up the element 24.

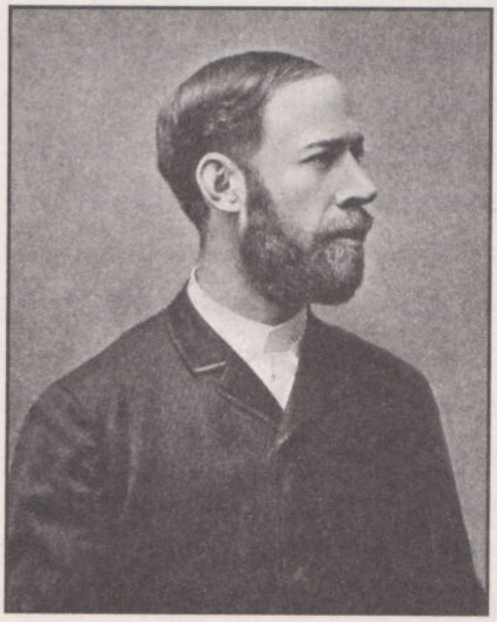
Consequently, the prime numbers are the primordial elements of which all numbers are built. The word 'prime' comes from the Latin *primus*, which means 'first' and includes the notions of 'primary' and 'primitive' in its original sense, becasue all numbers can be derived from them. In the same way that atoms combine to form molecules, prime numbers combine to form composite numbers. All the known chemical elements are composed of atoms that combine with each other in specific ways. Dmitri Ivanovich Mendeleyev (1834–1907) was the Russian chemist who created the periodic table of elements, an arrangement in which all the chemical elements are grouped. However, there is nothing analogous for prime numbers, no table enabling them to be grouped according to a rule, no law that generates them all without at least some anomalies. Prime numbers occur as a chaotic set, without rhyme or reason, and are distributed in an apparently random fashion throughout the series of natural numbers.

Prime numbers: invention or discovery?

As soon as a counting system has been established, it seems logical that the first thing to check is whether a number is odd or even. The next step is to consider factorising numbers; this establishes the criteria of division, which are taught at an early stage in schools. Thus, a culture that has established its counting system has a collection of numbers defined by just a few properties that are easy to determine. This does not apply to prime numbers. The only thing known for sure about these numbers is that they cannot be even - with the exception of 2, the only even prime - as they would then be divisible by 2. Neither would it be right to treat them as rarities that are hard to find, because Euclid demonstrated that they are infinite – we shall describe the elegant way in which this was proven later on. And it is not possible to underestimate their importance, as the fundamental theorem of arithmetic has given them a starring role in mathematics. Therefore, as has already been said, the primes have become an object of study in their own right.

When we speak of an object of scientific study, it seems reasonable to assume that the object concerned is out there somewhere. We may or may not have discovered it yet and may then investigate or ignore it, but in any event it continues to exist, whatever we think or do about it. At a certain point in history, bacteria became objects of study for biologists. Nobody doubts that they were present as living organisms before biologists existed, indeed a long time before the human species emerged. No one questions this in any scientific circles. However, in mathematics the issue takes on a different complexion. Are prime numbers an invention of the human mind or a discovery? Would prime numbers exist if humans did not? The matter has generated and continues to generate a lot of discussion – which is exciting for some and unimportant for others. Most likely it is a question without an answer and on which we can only express an opinion.

What is really important about the nature of the mathematical mind is that the researcher behaves like an explorer entering strange, unseen places, as if mathematics were actually separate from him. This sense of adventure is part of the very essence of mathematical research and is what gives it a poetic flourish. The German physicist Heinrich Rudolf Hertz (1857–1894) posed the question: "Is it possible not to feel that mathematical formulae have an independent existence and an intelligence of their own, that they are wiser then we are, wiser even than their discoverers, and that we obtain from them more than we originally put in?"



The universality of mathematics questions whether numbers have an independent existence beyond the human mind. This thought preoccupied the German physicist Heinrich Rudolf Hertz.

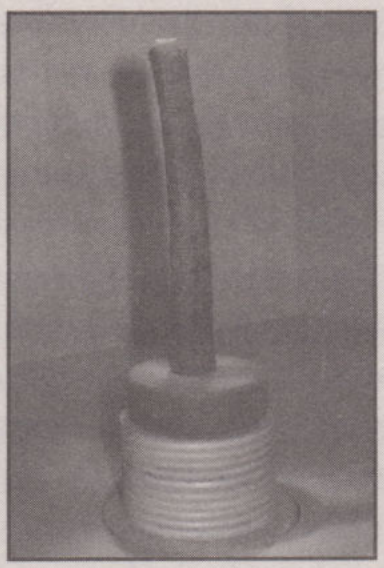
The philosophical or, perhaps rather, the epistemological school that accepts that objects (including mathematical truths) exist in their own right is known as Platonism. This states that an objective position can only be taken so long as one is in the presence of objects.

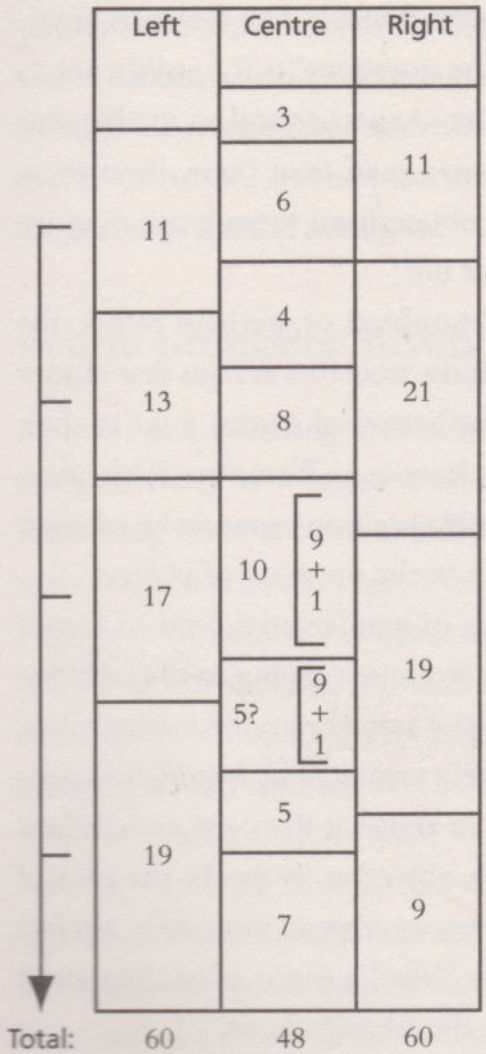
Historians of mathematics tend to favour the Platonist argument owing to the indisputable fact of the universality of mathematics: Cultures widely separated by history and geography tend to arrive at the same conclusions and the same objective truths. In the case of prime numbers, we have an interesting artefact that might be called a piece of mathematical archaeology: the Ishango bone.

THE ISHANGO BONE

The Ishango bone is probably a baboon's fibula and at first glance, it looks as if it has been carved into a tool of some kind. It appears to be a handle, easily gripped and with a sharpened crystal of quartz at one end. It was found near the sources of the Nile, on the border between Uganda and the Democratic Republic of Congo, and belonged to a tribal society that was buried by a volcanic eruption. The bone is thought to be around 20,000 years old.

The Ishango bone is displayed in the Museum of Natural Sciences in Brussels, Belgium.





The bone contains markings in the form of short straight lines. A detailed study has suggested that, rather than a tool, this is a numbering system to assist with counting. In that case, it is likely that the quartz tip was used to mark figures in some way. In other words, the bone handle could act as a primitive calculator. The distribution of marks in the columns suggests addition and multiplication operations in a base 12 counting system. The numbers on the right are all odd, but what is really surprising is that all the ones on the left are primes, specifically those falling between 10 and 20. It would seem very unlikely that these marks are distributed purely by chance or in any other way that does not suggest an advanced method of calculation. Let us recall that the

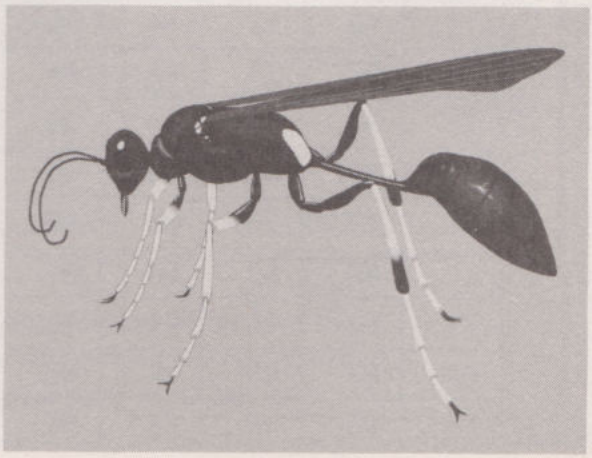
A diagrammatic version of the Ishango bone, showing the distribution of marks, in three columns. The bone seems to have been used to perform mathematical calculations.

concept of a prime number requires a process of abstract thought that goes beyond mere counting techniques.

The question of whether mathematical truths exist independently of human beings has a third answer, a compromise solution that allows for the possibility that there are indeed mathematical objects to be discovered, but they are 'mental objects' predetermined by our genetic heritage. If this is so, some primitive form of these notions should also exist in nature. Regarding the ability to count, there are several examples of species in the animal kingdom that can do so quite accurately. Solitary wasps can count the number of live caterpillars that they leave with their eggs as provisions for each larva to eat after they hatch – it is always precisely 5, 12 or 24. Among the species belonging to the genus *Eumenes* we find an even more astonishing example. The wasp knows whether a male or female will emerge from the egg it has laid. Putting aside how it manages to ascertain the sex of its offspring, as the chambers in which it lays the eggs and deposits food are apparently identical, the wasp leaves five caterpillars per male egg and ten for every female one. The reason for this difference is that female wasps grow much larger than males.

Even for a more complex notion, such as prime numbers, there is a curious example involving a few species of so-called periodical cicadas, including *Magicicada septendecim* and *Magicicada tredecim*. The species names *septendecim* and *tredecim* mean 17 and 13, respectively, and refer to the length of the life cycles of both insects. Both are prime numbers, and zoologists have formulated different theories to explain the choice of a prime number of years for the life cycles of these insects.

Let us take the example of *Magicicada septendecim*. This cicada lives as a nymph underground and feeds on sap, which it sucks from tree roots. It spends 17 years in this state and then emerges on the surface to become an adult insect, a stage that lasts only a few days during which time it breeds and finally dies. The theory



The females of some solitary wasps lay eggs in chambers that are also provisioned with several paralysed caterpillars that will provide the first meal for the wasp larvae after they hatch. The astonishing thing is that these wasps always leave the same number of caterpillars and know whether the egg concerned will hatch as a male or a female, as this also dictates the number of 'meals' the mother will leave for her young.

explaining this behaviour is as follows. The adult cicada is exposed to a parasite with a life cycle that is two years long. If the cicada's life cycle were a multiple of 2, both species would coincide every two, four, eight years and so on. The same would happen with any other multiple. However, if the life cycle is a sufficiently large prime number of years, for example 17, the parasite and the cicada can only coincide every 34 years, the first multiple of 17. If, for the sake of argument, the parasite's life cycle were 16 years, they would only meet once every $16 \cdot 17 = 272$ years.

It is quite likely that, over time, the study of animal behaviour will yield more examples of species that can count. We should not make too much of the simplicity of these examples, but the important fact remains that, although mathematical objects such as prime numbers are a human creation, the best researchers can experience them as if they exist independently.

Eratosthenes' sieve

Generating prime numbers has always been a thorny issue. One of the first known methods devoted to this activity is attributed to Eratosthenes of Cyrene (273–194 BC), a Greek mathematician, astronomer and geographer, who was also director of the Library at Alexandria. The method is known as Eratosthenes' sieve. Let's see how the sieve produces the primes within the first 100 natural numbers.

First, we draw up a table with all the natural numbers from 1 to 100. We start by eliminating all numbers that are multiples of two: 4, 6, 8, 10…; then all those that are multiples of three: 6 (already eliminated), 9, 12, 15, … The same is also done for multiples of five and seven:

~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	53	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59	~~60~~
61	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	67	~~68~~	~~69~~	~~70~~
71	~~72~~	73	~~74~~	~~75~~	~~76~~	~~77~~	~~78~~	79	~~80~~
~~81~~	~~82~~	83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89	~~90~~
~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	97	~~98~~	~~99~~	~~100~~

The numbers that are left are primes.

Note that the 'sieving' ends when the number 10, the square root of 100, is reached. In general, to find all primes smaller than a given number N, all numbers less than or equal to $\sqrt{N}$ should be 'sieved'. This provides a method of finding primes that are smaller than a particular number. The method continues to be used today, more than 2,000 years after it was invented, for finding 'small primes', so called because they are under ten billion.

DIMENSIONS OF THE EARTH

Eratosthenes is associated with the prime-number sieve that bears his name. However, this was by no means his most important achievement. He was not only a talented mathematician, but also a talented geographer and astronomer. In fact, Eratosthenes made many contributions to the progress of science. These include making a surprisingly accurate measurement of the circumference of the Earth. Using the techniques available in the 3rd century BC, he was able to calculate the polar circumference with an error margin of less than 1%, a remarkable accomplishment.

Map showing the world as it was known to Eratosthenes. The Greek scholar was the first to divide the planet into regular parallel sections, although his meridians were spaced irregularly.

How many prime numbers are there?

If we want to start thinking about the nature of prime numbers in order to find a relationship connecting them or a rule allowing us to predict where the next one will fall, we first need a fairly large set to investigate. The following list, obtained by using Eratosthenes' sieve, contains the prime numbers occurring in the first 1,000 natural numbers:

2	3	5	7	11	13	17	19
23	29	31	37	41	43	47	53
59	61	67	71	73	79	83	89
97	101	103	107	109	113	127	131
137	139	149	151	157	163	167	173
179	181	191	193	197	199	211	223
227	229	233	239	241	251	257	263
269	271	277	281	283	293	307	311
313	317	331	337	347	349	353	359
367	373	379	383	389	397	401	409
419	421	431	433	439	443	449	457
461	463	467	479	487	491	499	503
509	521	523	541	547	557	563	569
571	577	587	593	599	601	607	613
617	619	631	641	643	647	653	659
661	673	677	683	691	701	709	719
727	733	739	743	751	757	761	769
773	787	797	809	811	821	823	827
829	839	853	857	859	863	877	881
883	887	907	911	919	929	937	941
947	953	967	971	977	983	991	997

A cursory glance reveals that prime numbers are completely unpredictable. For example, there are more primes between 1 and 100 than between 101 and 200. Between 1 and 1,000 there are 168 primes. We might suppose that if our table were much bigger we would see that the number of primes increases as we move through the thousands. But this is not the case. Huge tables now exist and it is known, for example, that in the 1,000 numerals between 10^{100} and $10^{100} + 1{,}000$ there are only two prime numbers. And these numbers are more than 100 digits long!

It seems that, in order to find a pattern, the best thing would be to have a table that contains all the prime numbers. All of them? What if there are very many? Never mind, with the methods that we have at our disposal today it is possible to run them through all kinds of sieves and tests, enabling a pattern to be found. It is clear that, when we are dealing with finite sets, however large they may be, a pattern will finally emerge, or at least one that will fit can be invented. However, the situation changes radically when we are dealing with infinite sets, and so we need to decide whether the number of primes is infinite or not. This is a question also raised by Euclid. His method of resolving it is so ingenious, so elegant and so straightforward that it is worth looking at in some detail.

Let us start with a short list of consecutive prime numbers, for example: 2, 3, 5.

We then multiply them together:

$$2 \cdot 3 \cdot 5 = 30$$

Now we add 1 to the result:

$$2 \cdot 3 \cdot 5 + 1 = 30 + 1 = 31.$$

It is clear that 31 divided by any of the prime numbers in the list 2, 3, 5 will yield a remainder of 1:

$$31/2 = 15 \text{ remainder } 1; 15 \cdot 2 + 1 = 31$$
$$31/3 = 10 \text{ remainder } 1; 10 \cdot 3 + 1 = 31$$
$$31/5 = 6 \text{ remainder } 1; 5 \cdot 6 + 1 = 31$$

This shows that the number 31 is not divisible by any of them. This always happens; if we start with a list of consecutive prime numbers, multiply them together and add 1 then the number obtained is not divisible by any in the list. This simple fact is at the heart of Euclid's demonstration.

The number 31 is a prime number not in the original list, which is therefore incomplete. Let us take the following list as an example:

$$\{2, 3, 5, 7, 11, 13\}.$$

We multiply them together and add:

$$2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30{,}030 + 1 = 30{,}031.$$

This is not a prime number, as it is the product of two other numbers:

$$30{,}031 = 59 \cdot 509.$$

Euclid had already shown that any natural number can be reduced to a unique product of prime factors. If we apply this result to the number 30,031, which is a composite number, it is clear that the primes in list $\{2, 3, 5, 7, 11, 13\}$ are insufficient to reduce it to factors, as prime numbers are missing from it.

The conclusion is as follows. However long the original list of prime numbers, when they are multiplied together and 1 is added, the result is a new number of one of the following two types:

1) It is a prime number not in the list.
2) It is a composite number that breaks down into prime factors not in the list.

Therefore, the list is always incomplete unless it is infinitely long.

Unfortunately, this is not a method for obtaining all the prime numbers, although it is a very important starting point as it indicates the extent of the problem and provides a perspective without which it would be impossible to design any strategy to solve it. We might not think it very important to demonstrate that there is an infinite number of primes, as this is an intuitive fact. However, we have to be very careful with prime numbers, as they are so 'rare' that they might end at any moment. Nevertheless, Euclid's theorem shows conclusively that this will not happen.

Chapter 2

Prime Numbers: the Elusive Rule

As we have already said, prime numbers are one of those important topics that takes us back to the very origins of mathematics and then leads us, along a path of increasing complexity, to the cutting edge of modern science. Hence it is a very useful thread to follow as we unravel the fascinating and intricate history of the study of primes, especially in terms of how it has developed, that is to say, how the set of accepted truths of which it is now composed has been assembled. In this chapter we shall see how successive generations of mathematicians have scrutinised the natural numbers in search of a rule that predicts the occurrence of primes – a rule that has just become more and more elusive in the process. We shall also examine in greater detail the historical context in which these people worked and to what extent their work involved mystical and quasi-religious practices in a strange blend that scarcely resembles the scientific methods applied today. Nevertheless, slowly and laboriously, the way was prepared for new paradigms of a type that would inspire Fermat and Euler in the 17th and 18th centuries and which are discussed in detail in the next chapter.

Genius in context

As with the history of science in general, several individuals are associated with great discoveries in the study of prime numbers. But these individuals would not exist without a rich legacy built by others to support them; geniuses do not emerge from nowhere. Hence we must not ignore the paradigms upon which such a legacy was constructed, and the cultural background that helped to ensure that scientific advances were made.

In the 1930s specialist bookshops began to sell a series of mathematics textbooks by Nicolas Bourbaki, a previously unknown author. It was a collection of texts that achieved some success in the mathematical community. Among other things it provided students with a good treatment of mathematical analysis that had

THE MATHEMATICAL GENERAL

Where did the name Bourbaki come from? According to the version of one of his most distinguished followers, André Weil, it can be traced to an anecdote from his student days. It appears that Cartan and Weil, among others, attended a class given by an obscure mathematician with a vaguely Nordic name, an indefinable accent and an extraordinary appearance, who presented a theorem by Bourbaki that was as astonishing as it was incredible and which was attributed to a French soldier, Denis Bourbaki (1816–1897), a famous figure in the Franco-Prussian War. The class turned out to be a practical joke played by a student, Raoul Husson, but Cartan and Weil found inspiration in the name of the general – his Greek heritage made it the perfect pseudonym under which to present a 'Euclidian reconstruction' of mathematics. And so Bourbaki became a great mathematician after all.

General Denis Bourbaki, inspiration of patriots and mathematicians.

not existed until then. However, the aim was not only to provide the market with new textbooks but also to unify certain sectors of mathematics, such as algebra and analysis, where a degree of chaos ruled thanks to the huge quantity of new results that had been obtained in recent years. Many people were surprised to discover that a mathematician called Nicolas Bourbaki had never existed, and that the name had been chosen to represent a small group of mathematicians, among them Henri Cartan (1904–2008) and André Weil (1906–1998), who wished to reconstruct mathematics, inspired by philanthropic motives. For a mathematical collective, the work of the Bourbaki Group is well documented, no doubt because it was a recent phenomenon. The same cannot be said of any other groups from antiquity, such as the schools of Pythagoras and Euclid, the work of which may now be attributed to a single person, but many scholars believe is just as likely to have been collaborations of several people.

Information hubs

It is a remarkable fact that advances in scientific knowledge in general and mathematical knowledge in particular are never down to one individual. It is true that some individuals are responsible for great discoveries, but they themselves are products of a mathematical community. This requires there to be journals, colleges and conferences where information can be gathered and communication networks established among scientists. Nowadays, of course, communications have achieved an unprecedented peak of efficiency. Online communication places a discovery or scientific advance within the reach of anyone who wishes to access it, whenever they wish to look. However, the need to store information for others to be able to use is common to all ages; it is a cultural bond uniting a society. In this respect, prime numbers are an unusual object of research. They are the focus of an endeavour that began at the dawn of history and which is not yet complete. Following its trail not only provides information on their mathematical nature but also helps to develop those points of contact, which, using modern terminology, we might call 'information hubs'. The Library of Alexandria is a classic example.

Alexandria

Ptolemy I, also known as Soter, became the first ruler of Alexandria. Attracting the best architects in the world, the city became an architectural wonder. A long causeway connected it to Pharos Island on which a lighthouse was built that would be a marker for Mediterranean mariners for 1,000 years. Then a library was founded, the fame of which has persisted throughout history. A lighthouse and a library made Alexandria the most important information hub of the ancient world, an aim that Ptolemy was intent on achieving at any cost. His first move was to recall Demetrius from exile, a tyrant whom Cassander, one of Alexander the Great's three heirs, had appointed governor of Athens. It was Demetrius who had kept the Lyceum founded by Aristotle operating. Despite having participated in power politics, Demetrius's true vocation was knowledge, and so he was delighted to receive Ptolemy's invitation to found a library in Alexandria that would gather and catalogue all the knowledge of the civilised world in a single place.

The harbour of Alexandria consisted of small islands protected by dikes with a single outlet to the sea, a great channel along which ships would enter and exit. Protection from naval attack was almost total. One of the most important districts

was the Brucheion, situated right in the centre of the city and containing the most important public buildings, including one, the Museum, dedicated to the Muses and devoted to music and the sciences, that is to say, melody, rhythm and numbers. When Demetrius learned that this centre of knowledge was patronised by one of the most powerful rulers of the known world, he did not hesitate to become its director. The first thing he did was to ask Athens to lend him texts by the most important thinkers and writers produced by Hellenic culture until that time. He had them copied, returned the copies to Athens and stored the originals with the texts that Ptolemy had acquired as spoils of war during his campaigns. His method of increasing the collection turned out to be very effective, if rather unorthodox. All original documents on ships putting into Alexandria's harbour were requisitioned and copied, the originals were placed in the library and the copies returned to the ships. This is how the so-called 'ship library' came into being. But those who held power and wealth in the Mediterranean soon learned of the ruse and quickly refused to comply. Demetrius then offered merchants an incentive. If they wished to trade at the important markets in the port of Alexandria, they should bring along manuscripts from their ports of origin as a way of ensuring safe passage; it did not matter whether they dealt with engineering, philosophy, art, mathematics or music, provided that they contributed to knowledge. The proviso was that copies would be made, the originals would stay in the library, and the copies would be returned to the merchants. These copies were placed in their original holders so that most owners did not even notice the difference, and, even if they did, it seems most were not unduly concerned. It is recorded that, at that time, Alexandria employed the largest group of scribes ever assembled.

But Alexandria was not just a centre where information was stored, it was also a place where it was processed. It quickly attracted many specialists in all disciplines, who gave classes and shared their knowledge with other scholars. Hence classrooms, accommodation, arcades and promenades were built for this purpose.

It is reasonable to suppose that a number of schools were established, including, the school of Euclid, which, like the Bourbaki Group, two millennia later, collected the mathematical knowledge of its time and turned it into a school of thought, that is, a way of thinking about and doing mathematics, the results of which are still valid today.

Let us remember that 2,000 years later, we are still teaching in schools precisely the same geometry that was born in the classrooms and gardens of ancient Alexandria.

Alexandria was the most important information hub in the ancient world. Above: An engraving depicting scholars at work inside the famous library. Left: Roman coins stamped with the image of the Pharos lighthouse, another of the city's wonders.

Large gaps

One of the first things about prime numbers to attract the attention of ancient mathematicians was the absence of rules predicting their occurrence in the sequence of natural numbers. And not only that, but their absence – the way in which they fail to occur – is just as erratic. Thus they may turn up relatively close together or, on the contrary, occur very far apart. For example, if we list the prime numbers that occur in the first 100 natural numbers:

2, 3, 5, 7, 11, 13, 17, 19, 23, 29,
31, 37, 41, 43, 47, 53, 59, 61,
67, 71, 73, 79, 83, 89, 97,

we see that the first eight primes appear quickly within the first 20 numbers, whereas there are none between 89 and 97.

If we set out the prime numbers between 100 and 200:

101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163,
167, 173, 179, 181, 191, 193, 197, 199,

we see large gaps, such as the series of nine composite (non-prime) numbers between 182 and 190.

So the following question arises. How is it possible that there are very large gaps, such as 50,000 consecutive numbers, in which not a single prime number occurs?

The set of prime numbers is big enough for large gaps to occur within it, with sequences of consecutive numbers of any length existing without a prime number. This conclusion is not mere conjecture but derives from a result that is easy to demonstrate.

Consider the product of the first four natural numbers:

$$1 \cdot 2 \cdot 3 \cdot 4.$$

We can be certain that the number $1 \cdot 2 \cdot 3 \cdot 4 + 2$ cannot be prime, as it is divisible

by 2. This can be immediately confirmed, as $1 \cdot 2 \cdot 3 \cdot 4 + 2 = 26$, and we divide it by 2 to obtain 13.

We did not need to perform any operation to know that it was divisible by 2, as the two numbers to be added together contained the number 2.

For the same reason we can say that:

$1 \cdot 2 \cdot 3 \cdot 4 + 3$ cannot be a prime as it is divisible by 3;
$1 \cdot 2 \cdot 3 \cdot 4 + 4$ cannot be a prime as it is divisible by 4.

Thus we have obtained three consecutive numbers, 26, 27 and 28, which are not primes. If we now wish to obtain four consecutive numbers that are not primes, we execute the following:

$$1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 + 2 = 122$$
$$1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 + 3 = 123$$
$$1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 + 4 = 124$$
$$1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 + 5 = 125.$$

It is more convenient to represent the product of consecutive numbers with an exclamation mark:

$$1 \cdot 2 \cdot 3 \cdot 4 = 4!$$
$$1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 5!$$

In mathematics, this kind of expression is called a 'factorial'. For example, the factorial of 6 is:

$$6! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720.$$

Therefore, it is more convenient to write the above expressions as follows:

$$5! + 2$$
$$5! + 3$$
$$5! + 4$$
$$5! + 5.$$

Thus we can write a series of consecutive numbers that do not contain any primes. For example, if we wish to write 100 consecutive numbers so that none of them is a prime, we have only to write:

$$101! + 2$$
$$101! + 3$$
$$101! + 4$$

and so on up to $101! + 101$.

This means that there are large gaps in which no prime numbers can appear. In the same way we could construct a series of five trillion consecutive numbers in which no prime number appears. This suggests that prime numbers become ever scarcer as we progress through the series of natural numbers and, therefore, as we approach infinity there will come a point when no more occur.

This tentative idea is based on a false premise, as we know that Euclid's fundamental statement in number theory states that there are infinitely many prime numbers and that, however long a series of composite numbers is, eventually a prime number will occur.

USING A CALCULATOR

It is tempting to imagine programs that would help to identify large gaps between prime numbers using computer power. Indeed, the algorithm would be fairly simple, but we should bear in mind that when we handle factorial expressions we can soon forget about the seemingly endless capacity of calculators. Factorials grow at a dizzying pace. You can check on any pocket calculator with a factorial key that this is so – remember that the symbol is !. The following is obtained with just the first few numbers:

$1! = 1$; $2! = 2$; $3! = 6$; $4! = 24$; $5! = 120$; $6! = 720$;
$7! = 5{,}040$; $8! = 40{,}320$; $9! = 362{,}880$;
$10! = 3{,}628{,}800$.

Most calculators cannot perform this function beyond the number 70.

A sense of rhythm

There is a point that may occur at some music concerts when the audience becomes animated and applauds in time to the music. Initially this appears to work, but after a while the synchrony between the rhythm of the audience and that maintained by the percussionist begins to be lost. The situation can be kept more or less stable in the case of simple rhythms, but it becomes unimaginable in the case of more complicated rhythms. We can use this analogy to understand the efforts of mathematicians to impose a sense of rhythm on prime numbers, along the lines of '1, 2, 3 ... go!'. It doesn't work; prime numbers do not occur after every three composite numbers. Let's try something else: '1, 2, 3, 20, 100... go!'. That doesn't work either. We could go on trying this *ad infinitum*. Even today we do not know whether this set of numbers has a devilishly complicated rhythm or if it simply lacks a sense of rhythm altogether.

How can we impose a pattern on a sequence of numbers? There are many ways of doing so. The important thing to note is that once it is achieved, then it should be possible to predict the next number in a sequence.

For example, the sequence:

$$2, 4, 6, 8, \ldots$$

is straightforward, as everybody knows that the next number is 10.

In the case of:

$$1, 3, 5, 7, \ldots$$

it is also easy to tell that the next number is 9. The first is a sequence of even numbers and the second a sequence of odd numbers.

Another example is:

$$2, 3, 5, 9, 17, \ldots$$

Here each number is obtained by multiplying the previous one by 2 and subtracting 1 from the result.

In mathematics, the pattern is revealed when we obtain what is called 'a general term', an expression that gives the value of each sequence entry by simply attributing values to n.

For example, in the sequence of even numbers, the general term is:

$$a_n = 2n$$

$$\begin{aligned} &\text{If } n = 1 \quad a_1 = 2 \cdot 1 = 2 \\ &\text{If } n = 2 \quad a_2 = 2 \cdot 2 = 4 \\ &\text{If } n = 3 \quad a_3 = 2 \cdot 3 = 6. \end{aligned}$$

In the case of the odd-number sequence, the general term would be given by:

$$a_n = 2n + 1.$$

We can use this system to find the value of any term. If we wish to find the value of the term occupying the 27th position in the sequence we need only to set $n = 27$ in the expression for the general term:

$$a_{27} = 2 \cdot 27 + 1 = 55.$$

Finding the formula for the general term is equivalent to discovering the law governing the sequence. The question then is, as we can obtain any entry in the sequence using the general term, can we find a general term from enough entries in a sequence? For many sequences the answer to the second half of that question is often a significantly more complicated problem.

For example, the number sequence:

$$\frac{2}{4}, \frac{5}{7}, \frac{10}{12}, \cdots$$

may not be so easy to predict, and indeed the general term of this sequence is:

$$a_n = \frac{n^2 + 1}{n^2 + 3}.$$

To find the first three terms, we assign the appropriate values to n:

$$a_1 = \frac{1^2+1}{1^2+3} = \frac{2}{4};$$
$$a_2 = \frac{2^2+1}{2^2+3} = \frac{5}{7};$$
$$a_3 = \frac{3^2+1}{3^2+3} = \frac{10}{12}.$$

This represents much of the effort made by mathematicians throughout history in studying prime numbers – an attempt to find any kind of rule always ending in frustrations and failures of every kind. How is it possible that this chaotic collection of numbers is governed only by rules of chance? In any case, mathematicians tend to qualify what they mean by failures, because if their studies fail, their research, which may not have attained their goals, will perhaps have blazed new trails, invented other ways of doing maths and opened doors to new paradigms. It often seems that the objective being sought was merely a pretext for tackling new problems. Hence, prime numbers have been and continue to be one of the most fruitful sources of paradox and conjecture.

Twin primes

Even if a general law cannot be established, it is at least possible to study the behaviour of some prime numbers with special characteristics. It is like standing in front of a door through which groups of people endlessly pass. We know that some are men and others women but we cannot devise a rule allowing us to predict when one group or the other will appear. Then one day we notice something peculiar and realise that men wearing hats, women with glasses and children with umbrellas occasionally appear. We then try to find a rule that defines the appearance of specific groups – for example, men with hats appearing 100 times as often as women or if every time one man enters he is always followed by a woman. This would allow us to establish some kind of pattern. We may indeed find one that works, only to discover that it fails when we have monitored the arrival of three million people. Then we would say: 'Oh, nearly!' and our research would be expressed by the words: 'it's almost as if ...', an expression that has been used often in the history of prime numbers.

THE SOLITUDE OF PRIME NUMBERS

Prime numbers can be separated by millions and millions of numbers or by just one, which is the closest they can ever get to each other; except for 2 and 3, primes are never adjacent to each other. This fact was used as a metaphor in the title of a recent work of literature, *The Solitude of Prime Numbers* by Paolo Giordano. The novel is about two lonely children, Mattia and Alice, both of whom had exposure to traumatic events as children, and who grow up to form a special relationship. In a paragraph of the novel the metaphor is stated explicitly: "In a class in the first year, Mattia learned that among prime numbers some are rather special. Mathematicians call these numbers twins. They are pairs of prime numbers that occur together or, rather, almost together, as they are always separated by a number that prevents them from touching completely: Numbers like 11 and 13, 17 and 19, or 41 and 43. Mattia thought that he and Alice were like that, twin primes, alone and lost, together but not close enough to be really in contact."

It is true that some families of prime numbers have been successfully characterised (several dozen in fact), and these have enabled some progress to be made over time. Here we will focus on some unusual pairs of prime numbers with characteristics that help us to understand a little better the mathematical difficulties posed by these most erratic of numbers

No two prime numbers can be consecutive, because every prime number is odd and the next number has to be even and hence not a prime. Therefore, two prime numbers will always be separated by at least one other number. The exception is 2 and 3, which are consecutive; furthermore, 2 is the only even prime.

Among the first 100 natural numbers we find the following pairs separated by two units:

$$(3, 5), (5, 7), (11, 13), (17, 19), (29, 31), (41, 43), (59, 61) \text{ and } (71, 73).$$

These pairs are called 'twin primes' or just 'twins'.

Twins can be described by the expression $(p, p + 2)$, where p is a prime number. Here is a list of all twin primes in the first 1,000 numbers:

(3, 5),	(5, 7),	(11, 13),	(17, 19),	(29, 31),
(41, 43),	(59, 61),	(71, 73),	(101, 103),	(107, 109),
(137, 139),	(149,151),	(179, 181),	(191, 193),	(197, 199),
(227, 229),	(239, 241),	(269, 271),	(281, 283),	(311, 313),
(347, 349),	(419, 421),	(431, 433),	(461, 463),	(521, 523),
(569, 571),	(599, 601),	(617, 619),	(641, 643),	(659, 661),
(809, 811),	(821, 823),	(827, 829),	(857, 859),	(881, 883).

We know that twin primes become scarcer as we progress through the sequence of natural numbers. However, thanks to computer analysis it is now clear that primes of this type continue to occur among extraordinarily large numbers, and this has led mathematicians to conjecture that there is an infinity of twin primes, just as there are infinite primes, but no one has yet been able to prove it. More recently, mathematician Yitang Zhang of the University of New Hampshire in Durham, has succeeded in proving a boundary: that there are infinitely many pairs of primes that are less than 70 million units apart.

Another conspicuous group of prime numbers among the first 100 natural numbers is that consisting of the numbers 3, 5 and 7. If p is a prime number, these three numbers can be expressed by $(p, p+2, p+4)$. Groups of this type could be called 'triplets'. In fact, there is no need to name them at all, as only these three exist. This is a confirmed result. Luckily, the matter is closed as otherwise the triplets would have generated another set of conjectures that would still remain unresolved.

The largest twin primes known (discovered in 2009) are those formed by the numbers $65{,}516{,}468{,}355 \cdot 2^{333333} - 1$ and $65{,}516{,}468{,}355 \cdot 2^{333333} + 1$, which consist of 100,355 digits!

INFINITE SEPARATIONS

Twin primes have given rise to a number of conjectures, in addition to the one stating that they are infinite. One of them is of a more general nature and was formulated in 1849 by the French mathematician Alphonse de Polignac (1817–1890). He suggested that for every number C, there are infinite pairs of prime numbers separated by $2C$ composite numbers. That is, there is an infinity of prime numbers separated by four composite numbers, by six composite numbers, by eight composite numbers and so on. For example, $C = 1$ represents the twin primes conjecture.

Magic and mathematics

We have emphasised the importance that information hubs have, and have had, throughout history. We now focus on another aspect that is of some significance for the history of mathematics, especially in terms of numbers. That is the possible relationship between magic and mathematics. By magic we mean a historical tradition of mathematics called arithmology or, more commonly, numerology. The relationship between mathematics and numerology is similar to that between astronomy and astrology or between chemistry and alchemy. Nowadays, these pairs have become practically separate, but throughout history they formed marriages of convenience that cannot be ignored if we wish to have an historical perspective on what occurred at each stage in the development of the field.

Numbers and hence prime numbers have not only been subject to mathematical research but also to philosophical investigation, and especially caught the attention of religious cults. When numbers become part of such a framework, they are used in very different ways. We find them in the Bible, in magic squares, in magic sums and, especially, in the philosophical foundation of the Pythagorean school for which geometrical figures and numbers were the basis of all existence.

Therefore, we shall encounter mysteries and legends surrounding famous mathematicians, like Mersenne or Fermat, who are said to have known only very simple mathematical methods but attained goals that were beyond others. The historian Libri said that "Fermat knew things that we do not, and to equal him we require methods more perfect than those invented later." We should not forget that, unlike many mathematicians of his period, Fermat was not one of those scientists who systematically hide their knowledge, even if he did conceal how he obtained it.

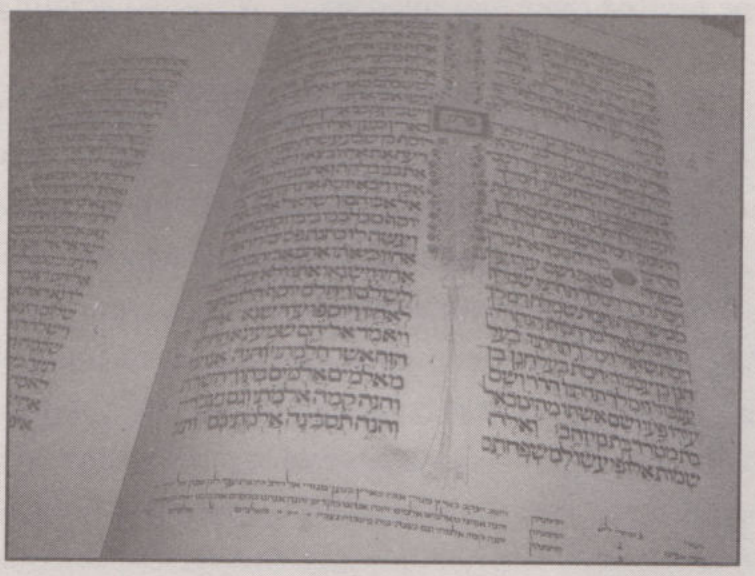
We will be visiting periods when mathematical rigour, as it came to be conceived in the 18th century, did not have the importance we ascribe to it today. The aim in those days was to create a mathematical tool box for practical purposes rather than theoretical ones. Thus traditional teaching, with all the mystical symbolism with which it was imbued, was not an impediment but rather the opposite – a space in which the imagination could take flight.

Hence we have a rather erroneous idea of what mathematics is, because we also have an erroneous notion of how the great mathematicians did their work. Ignorance of what mathematicians do not only generates ignorance of the nature of the mathematical mind but, to some extent, it has also been the source of the subject's unpopularity. The final result of a piece of research, which usually takes the form of

a theorem, has been so arranged, revised and polished that it almost always appears somewhat obscure to people who have not had previous training. It is difficult to make someone understand the beauty that may reside in statements that are so technical and so utterly logical. However, the researcher has not journeyed in that way but roamed through a dense forest of numbers in which paths are scarcely visible and where it is frequently pitch dark.

THE BOOK OF NUMBERS

Numbers is the fourth book of the Bible and part of the five-book *Torah* attributed to Moses. At first glance, *Numbers* is a book of accounts and hence is of undoubted historical value, as it meticulously lists all the quantities it contains, from tribal chiefs to heads of cattle, so forming a historical background for the events reported elsewhere in the holy books. However, it is also a book of secret codes for those initiates who can decipher its messages, because the lists of numbers not only represent quantities but also have meaning. For example, 1 symbolises God, 2, man, 3, the totality of things, etc. It is curious that the number 5 represents an undefined quantity – 'several'. For example, in the multiplication of the loaves during the Sermon on the Mount, it is said that Jesus took five loaves, that is, 'some loaves'. The strangeness resides in the fact that 5 is the maximum number of objects we can identify with a single glance. It is known that we can recognise the size of sets of up to four objects without actually counting them; if there are more, we have to divide them into groups of four or less and add them together.



Torah *is known by Christians as the* Pentateuch *and is the first five books of the* Old Testament.

The fact that the mathematical mind explores this most obscure intellectual landscape has also troubled some guardians of morality. A good example of this are the words of Saint Augustine on the matter: "The good Christian must be on his guard against mathematicians and all who make empty prophecies. There is a risk that mathematicians have made a pact with the devil, and their task is to lead a man's spirit astray in order to send him to hell."

In addition to what we have called information hubs and the magical aspects of numbers, there is a third point that we should bear in mind when pursuing the course of prime numbers through history. This is the exceptional gift for numbers with which some people are blessed – an ability that in most cases goes hand in hand with a gift for words. Most of the famous mathematicians whom we shall see 'patrolling' prime numbers also possessed an extraordinary gift for languages. This in itself it not surprising, because, as we said at the beginning of the book, numbers and words are related as the most ubiquitous abstract concepts used by humans. In earlier periods, when devices to assist calculation were practically non-existent, an ability to do mental arithmetic accurately was an essential strength of great mathematicians. This ability went beyond mere numerical computation, which belongs to the world of the showman rather than that of the mathematician. Men of the stature of Fermat, Mersenne, Euler and Ramanujan had the magic gift for 'seeing' into the world of numbers. This ability enabled them to discover relationships that only they could appreciate – relationships requiring proofs that often remained beyond their reach and, in some cases, even beyond their best interests.

HUMAN CALCULATORS

Stage calculators appeared in the 19th century, performing arithmetic to entertain crowds. They soon become fashionable and put on shows in European and American theatres, attended by an audience fascinated by such astonishing mental feats. Zerah Colburn, the first of the professional calculators to be properly documented, was born in Cabot, Vermont (USA) in 1804. On one occasion he was asked to multiply 21,734 by 543. Almost immediately he replied 11,801,562. Someone in the audience asked him how he did it: "I saw that 534 is equal to three times 181. I first multiplied 21,734 by three and then the result by 181", replied Colburn who normally took just a few seconds to multiply five-digit numbers. This occurred in 1812 when Zerah Colburn was just eight years old.

Chapter 3

New Paradigms

The middle of the 17th century saw the rise of an important scientific movement that reached beyond traditional academic institutions. By then many European universities, centres dedicated to pushing back the frontiers of knowledge, were well established, but they had become somewhat resistant to new ways of learning. This was a serious problem for anyone wishing to pursue research outside the strictures of academic circles, because salaries were paid only to those working within the main centres of learning. Thus a period of patronage began, in which nobles and powerful landowners were proud to support great thinkers and be associated with the new ideas that were being proposed. Most biographies of the time cite not only the names of great scientists but also those of their patrons. However, this could create a bit of a communication problem at times.

Specialised institutions were established to meet the needs of scientific communication. As one example, what would become the future Académie des Sciences (founded by Colbert in 1666) started out in a friar's cell at a Parisian monastery. The man who lived there was Father Mersenne.

Marin Mersenne

Mersenne was born on 8 September 1588, in Oizé, in the present-day French department of Sarthe. There is very little information on the early years of his life. It is known that in 1604 he became a boarder at La Flèche, a school founded in 1603 by Henri IV and run by the Jesuit order, where he stayed for a year. During this time he became close friends with René Descartes, a fellow pupil. Their friendship would last throughout their lives.

Marin Mersenne (1588–1648).

In 1609 Mersenne began to study theology at

FRÈRES MINIMES: THE ORDER OF THE MINIMES BROTHERS

The name of this order reflects the fact that all its members were required to observe a minimum of religious principles. Its aim was to avoid any body of doctrine based on a set of revealed truths that would impose excessively strict rules of conduct. In fact, the only thing the members of the order were absolutely against was atheism. Essentially, they dedicated themselves to prayer, study and teaching, and endeavoured to ensure that their religious convictions never interfered with education or scientific development. Proof of this is the heated defence of the works of Galileo put up by Mersenne.

the Sorbonne, graduating two years later and entering the Frères Minimes order of monks. In 1612 he was appointed priest at the Monastery of the Annunciation in Paris. Between 1614 and 1618 he taught philosophy at Nevers Monastery. He returned to his cell at the Monastery of the Minimes Brothers where he remained until his death on 1 September 1648. Wanting to serve the aims of science to the end, Mersenne instructed in his will that his body be donated to the Faculty of Medicine for anatomical study.

Mersenne's first works were of a purely theological nature and include *Well-Known Questions on Genesis* (1623), *The Truth of Science against the Sceptics and Pyrrhonians* (1625), and *Theological, Physical, Moral and Mathematical Questions* (1634). His scientific works include *Universal Harmony* (1636), in which he devised a formula connecting the length of a string and the sound it emitted when plucked.

This formula allowed him to create a scale in which all intervals were equal, thereby dispensing with the then famous Pythagorean comma (the smallest difference between notes in Pythagorean tuning) and laying the foundations of what would become one of the greatest revolutions in the history of music, the chromatic or tempered scale.

Mersenne numbers

Mersenne's greatest scientific work of a purely mathematical nature was *Cogitata Physico-Mathematica* (1644) in which his famous treatment of prime numbers appears. In its introduction Mersenne states that among all the primes between 2 and 257, the number 2^p-1 is only prime if p is one of the following numbers:

$$2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257.$$

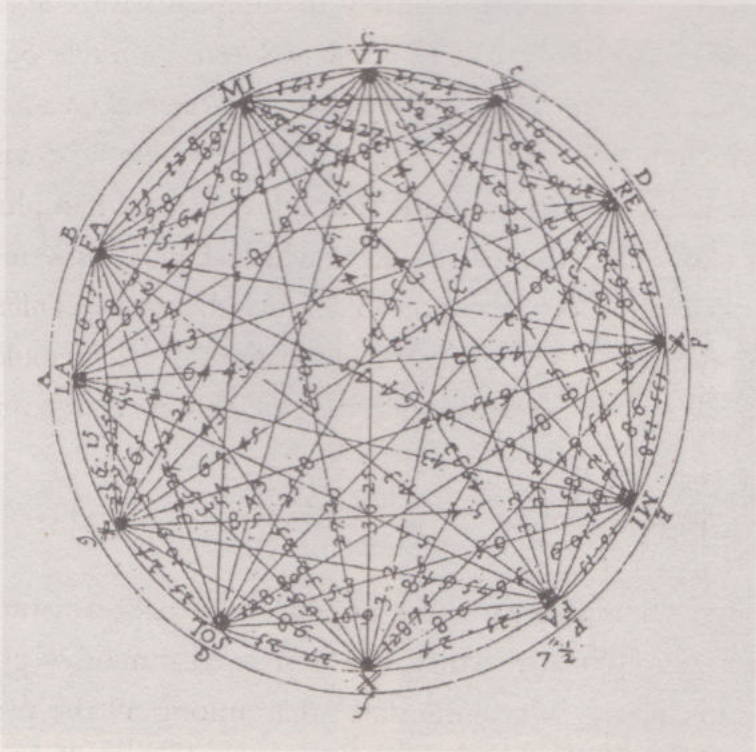
If we take the number 2 and raise it to the power of the last number in the list we get a number 77 digits long. How Mersenne managed to ascertain that this number is a prime with the means of calculation at his disposal is a mystery that no one has been able to solve.

It is easy to show that, if $2^p - 1$ is a prime, then p must be a prime (or, equivalently, that, if p is not a prime, then neither is $2^p - 1$). This result, which was already known in Mersenne's day, led him to investigate what would happen when a number p that is a prime is put into the expression. It was also known that $2^p - 1$ is a prime for the values $p = 2, 3, 5, 7, 13, 17$ and 19, but not for $p = 11$.

A hundred years would pass before Euler succeeded in showing that $2^{31} - 1$ is a prime. In 1947 the full list was finally produced, thus:

$$p = 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107 \text{ and } 127,$$

showing that Mersenne's original list had two incorrect numbers and was missing three more. Nevertheless, these numbers continue to be called 'Mersenne numbers', numbers that nowadays play an important role in so-called 'primality tests', which use a set of algorithms to determine whether or not a number is prime.



Mersenne studied vibrations in strings and created a scale divided into 12 equal intervals.

NERVE CENTRE

The small cell in which Mersenne spent the last 30 years of his life, at the Monastery of the Minimes Brothers next to the Palais Royal, became the nerve centre of European science. It came to be said that informing Mersenne of a discovery was equivalent to distributing a publication throughout Europe. After his death, documents were found in the cell showing that Mersenne maintained 78 lines of communication, in as many lines of research, with individuals from the world of science as notable as Torricelli, Decartes, Pascal, Gassendi, Roberval, Beaugrand and Fermat.

Pierre de Fermat

Fermat (1601–1665) has become a true legend in the world of mathematics. His discoveries, especially in the theory of numbers, a branch that he can be considered to have founded, have ensured his lasting fame as the 'prince of amateurs'. Furthermore, he had complete mastery of the classical languages, Latin and Greek, and of most of the European languages spoken at that time.

Fermat enjoyed wealth and privilege, which enabled him to indulge his passion for numbers to the full. He was born into a rich family and his legal studies won him a position as a local government official in Toulouse. One of the requirements of this public post was to refrain from all kinds of social activity in order to avoid any suspicion of corruption. He married Louise de Long, a cousin of his mother, and they had three children. The eldest, Climent-Samuel, who would later publish his father's work; Fermat's two daughters became nuns.

Fermat hardly ever travelled; the only journey he made was to Paris where, with an introduction from Pierre de Carcavi (1600–1684), an influential French mathematician, he met Father Mersenne at his monastery.

Some people who are fond of growing flowers spend a lot of time cultivating new varieties, either from seeds from distant countries or by breeding hybrids that could potentially yield pleasant surprises. Fermat cultivated numbers. One morning he looked into the garden of his mind and encountered a new species, which to ordinary folk, appeared to emerge as if by magic. He was not one of those mathematicians who hide their results, as he offered them to everyone, but he hardly ever explained how he had arrived at them. The expression "any number of the form $4n + 1$ is the sum of two squares" was,

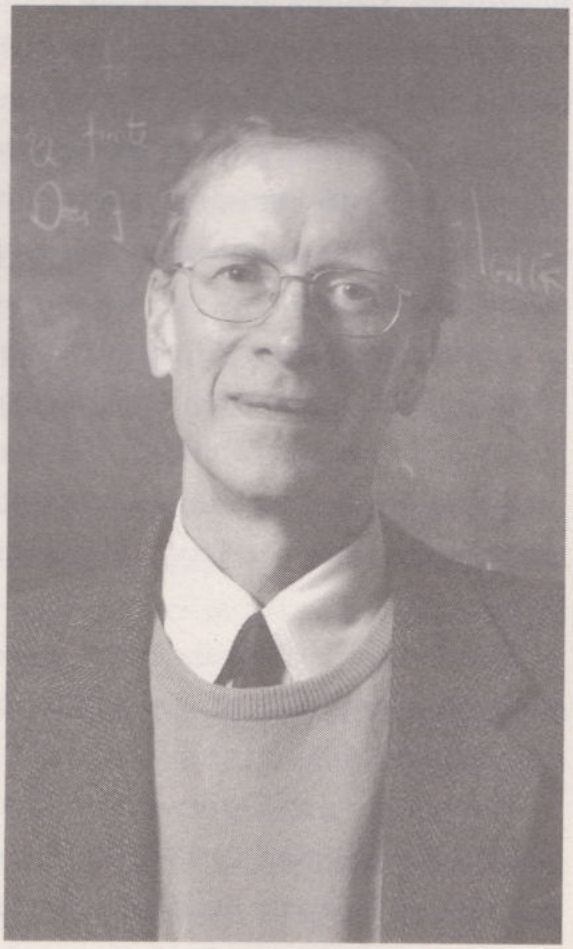
for example, one of the many results that Fermat never explained and was only proven by Euler in 1749 after seven years of hard work. Gauss commented that this result was "one of the most beautiful flowers that Fermat discovered in the garden of numbers."

Fermat's little theorem

In 1995 Andrew Wiles made Fermat front-page news when he proved one of the most famous conjectures in history: if n is an integer greater than 2 ($n > 2$), then there are no integers x, y and z, other than 0, that satisfy the equation:

$$x^n + y^n = z^n.$$

This is the conjecture known as 'Fermat's last theorem'.



Fermat's last theorem was solved in 1995 by the British mathematician Andrew John Wiles. Two years later he published an initial proof which, however, contained an error that he was later able to correct.

However, there is another theorem, much less well known, called 'Fermat's little theorem', which has turned out to be particularly relevant to the theory of prime numbers. It was first set out in a letter sent by Fermat on 18 October 1640, to Bernard Frénicle de Bessy (1605–1675), a friend and also an amateur mathematician, with whom Fermat shared some of his results (both were members of Mersenne's select circle). The letter stated:

> Every prime number is equivalent to a power minus one of any progression in which the index is a multiple of the given prime minus one... And this proposition is generally true for all progressions and all prime numbers; I would send you the proof if I didn't think it too long.

As usual, Fermat omits the proof, making the excuse that it was too long, as was also the case with his more famous final theorem. Most historians agree that it is more likely that the great man did not have a proof of these or many of the other conjectures he arrived at. In any event, Fermat considered himself an amateur mathematician and this allowed him to take certain liberties.

The statement appearing in the letter sent to Bessy is rather cryptic and muddled, and so it is expressed below in modern terminology.

It says that two numbers are related primes (or co-primes) when they have no factors in common.

For example, 8 and 27 are related primes, as they have no factors in common: $8 = 2^3$ and $27 = 3^3$. On the other hand, 12 and 15 are not co-primes, as they both have 3 as a common factor: $12 = 3 \cdot 4$ and $15 = 3 \cdot 5$.

Therefore, the theorem asserts that if p is a prime number and a any other number, for a and p to be co-primes, then it must be true that $a^p - a$ is divisible by p.

For example, let us take the prime number 3 and the number 8, which contains this prime: then $8^3 - 8 = 512 - 8 = 504$ is divisible by 3. In effect we are saying that $504/3 = 168$.

We can say that Fermat's little theorem is small but powerful (the adjective 'small' was first used in 1913 by the German mathematician Kurt Hensel), as it is one of the most frequently used theorems in primality tests to determine whether a very large number is prime.

In fact, Fermat himself must have used it as a mathematical tool to reduce large prime numbers to factor products. It is known, for example, that he was

THE CHINESE HYPOTHESIS

There are documented sources which suggest a possibility that 2,000 years before Fermat, Chinese mathematicians had formulated a system, known as the 'Chinese hypothesis', with a result very similar to that achieved by Fermat's little theorem. This hypothesis states that p is a prime number if and only if $2^p - 2$ is divisible by p. The Chinese hypothesis has hitherto been considered as a special case of Fermat's little theorem. However, the reverse, which states that if the condition is fulfilled then p is prime, is not true, and so the Chinese hypothesis although quite impressive, is considered only half right; as a whole it must be considered wrong.

able to find that 100,895,598,169 is a product of the numbers 898,423 and 112,303, both primes, in response to a question from Mersenne, who wished to know whether the larger number was prime. Even so, it is difficult to see how Fermat was able to manipulate such big numbers.

The theorem was first proved by Euler in 1736. (Leibniz had a similar proof but never published it.) Moreover, Gauss included another proof in his famous book *Disquisitiones Arithmeticae*, published in 1801. Euler would later come up with two further proofs. Of these, the simplest is Euler's first proof as it can be understood with a basic grasp of mathematics (see the Appendix).

Remember that Fermat's little theorem is a way of ascertaining whether or not a number is prime without having to find any of its factors. We can show this with a simple example.

Let us suppose that $p = 9$ and $a = 2$; then $2^9 - 2 = 510$, which is not divisible by 9, and so we conclude that 9 is not a prime number, something we already knew. The great usefulness of the simple method lies in the fact that it can be applied to very large numbers.

We should be aware that Fermat's little theorem posits an essential but insufficient condition, if p is prime, the condition is satisfied, but that alone does not mean that p has to be prime.

For example, if we take the number $p = 91$. It is a composite number as $p = 7 \cdot 13$. But, considering $a = 3$ easily shows that 91 divides into $3^{91} - 3$. Accordingly, composite 91 is the pair ($p = 91$, $a = 3$) verifies Fermat's little theorem.

The pair ($p = 341$, $a = 2$) provides a similar counter example.

Fermat numbers

A 'Fermat number' is a natural number of the following kind:

$$2^{2^n} + 1 .$$

It is usually indicated by the letter F (for Fermat) with a subscript (n) of the number concerned, so that F_0 is the first Fermat number, F_1 the second and so on. Let's calculate the values of the first five Fermat numbers – remember that any number raised to the power 0 equals 1:

$$2^0 = 1; 2^1 = 2; 2^2 = 4; 2^3 = 8.$$

Substituting into the above formula, we get:

$$F_0 = 2^{2^0} + 1 = 2^1 + 1 = 3$$
$$F_1 = 2^{2^1} + 1 = 2^2 + 1 = 4 + 1 = 5$$
$$F_2 = 2^{2^2} + 1 = 2^4 + 1 = 16 + 1 = 17$$
$$F_3 = 2^{2^3} + 1 = 2^8 + 1 = 256 + 1 = 257$$
$$F_4 = 2^{2^4} + 1 = 2^{16} + 1 = 65{,}536 + 1 = 65{,}537$$

Fermat conjectured that all numbers obtained in this way are prime. The first five numbers, 3, 5, 17, 257 and 65,537, are.

If $n = 5$, the number obtained is:

$$F_5 = 2^{2^5} + 1 = 2^{32} + 1 = 4{,}294{,}967{,}296 + 1 = 4{,}294{,}967{,}297.$$

Fermat was unable to say whether a number greater than four billion is prime. But Euler was, and in 1732 he managed to factorise this number as a product of two others:

$$4{,}294{,}967{,}297 = 641 \cdot 6{,}700{,}417.$$

Euler had caught Fermat making a false conjecture. It was the first time that anything like this had happened. Although the conjecture was false, Fermat numbers have been very productive, not only because they have obviously raised new questions and conjectures, but also because they have turned out to be useful for developing primality tests.

At present it is known that only the first five Fermat numbers are prime, which does not mean there are no more – in fact there may be an infinity of them. Full factorisation has only been done up to $n = 11$. Factorising a number as a product of primes is not an easy task. As we shall see later, this difficulty forms the basis of one of the most popular encryption methods used today.

Leonhard Euler

There is no branch of classical mathematics, be it calculus, differential equations, analytical and differential geometry, number theory or series, in which the name of the Swiss mathematician and physicist Leonard Euler (1707–1783) does not appear. He was one of the most prolific mathematicians of his time. After his death in Saint Petersburg, his writings continued to be admired and were published year after year by the Saint Petersburg Academy of Sciences. His complete works, thought to occupy around 90 large volumes, have yet to be published in full by the Swiss Academy of Sciences.

A Swiss 10 franc bank note from 1997, with a portrait of Euler on the front and a hydraulic turbine, the Solar System and light passing through various lenses on the back. All of this alludes to Euler's contribution to mathematics.

Euler always had a special interest in prime numbers. He constructed tables of all primes between 1 and 100,000 and devised formulae that enabled him to obtain an astonishing number of them. One of the most interesting is:

$$x^2 + x + q,$$

yielding prime numbers for certain values of x, for example, for between 0 and $q - 2$ such as $q = 2, 3, 5, 7, 11$ and 17. This was experimental mathematics, the aim of which was to achieve practical results, and so rigorous proofs were often lacking. However, unlike Fermat, Euler did not conceal his workings. If he had a proof he published it – and if he didn't it was because he didn't have one.

Euler brought about a change in the world of mathematics, the result of a slow but inexorable shift in thinking. Among his many contributions, there are three that had a crucial impact on later research into prime numbers: the concept of a function, infinite sums and the use of imaginary quantities – we shall be returning to the last of these later on.

Functions

Euler established the firm foundations of what in later centuries would be called 'mathematical analysis'. It was he who introduced the notation we currently use to represent a function, $f(x)$. A function works as a device for transforming numbers into other numbers according to an established rule. (We refer here to real functions with real variables only.) For example, if the rule states that a fixed quantity, such as 3, should be added to the number in question, the function is written as follows:

$$f(x) = x + 3.$$

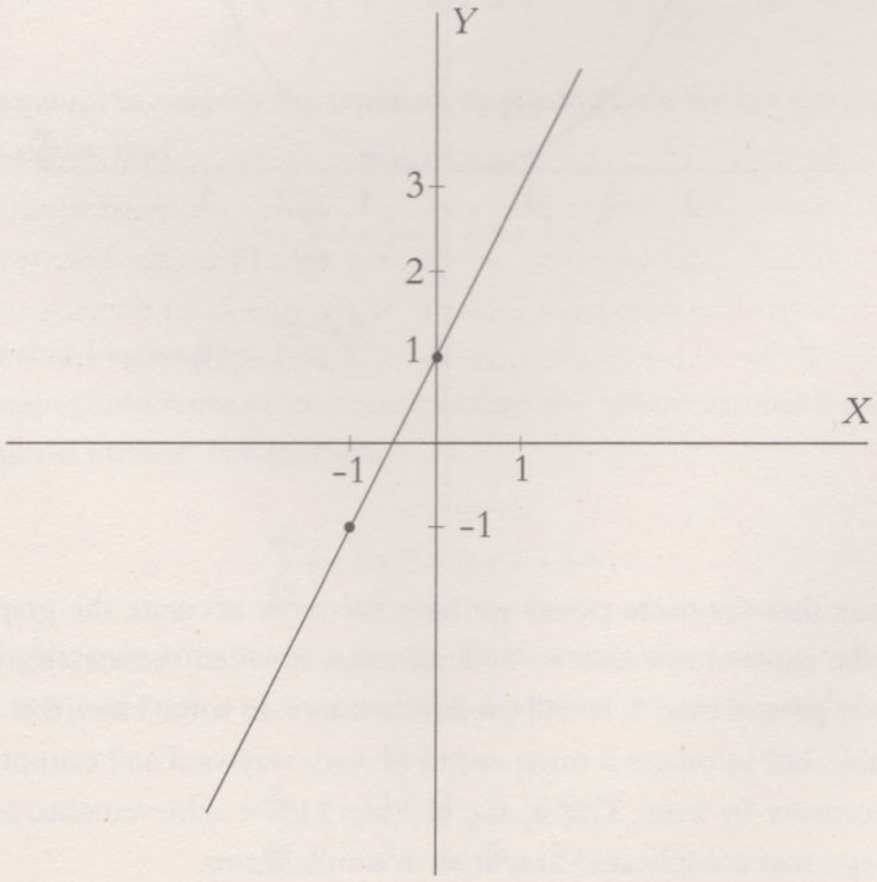
The function can now be applied:

$$f(1) = 1 + 3 = 4$$
$$f(2) = 2 + 3 = 5$$
$$f(24) = 24 + 3 = 27$$
$$f(0.32) = 0.32 + 3 = 3.32.$$

A real function with a real variable will assign each number with another real number. For example, the function $f(x) = 2x + 1$ assigns to each value of x twice its value plus 1. A simple table of values, such as this:

$f(x) = 2x + 1$	
x	$f(x)$
1	3
2	5
3	7
–1	–1
–2	–3
–3	–5

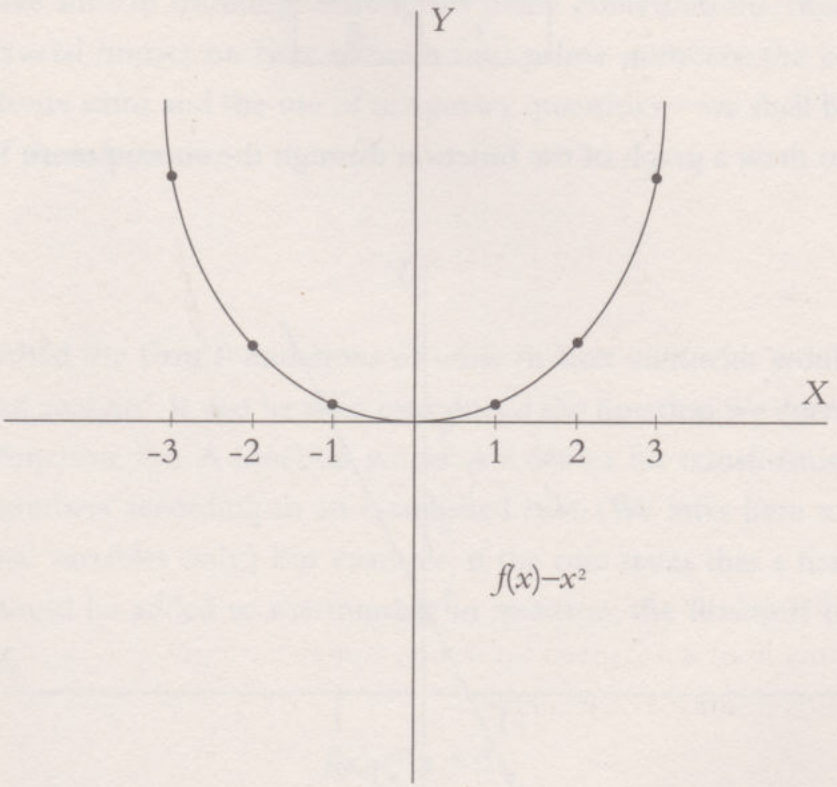
enables us to draw a graph of the function through the above points:



This is a very simple case, as it involves a straight line and so, to draw it, we need only two points. However, a function like $f(x) = x^2$, which would have a table like the following:

x	x^2
1	1
2	4
3	9
4	16
...	...

... is not so easy to draw:



It is a fact that the more points we have the more accurate the graph will be, but when the expression ceases to be linear, that is, when the variable x is raised to any power greater than 1, it will produce a curve. In some cases, this curve will be predictable, but in others it turns out to be very wayward and cannot be drawn with any accuracy by hand. This is one of Euler's great achievements, namely, the ability to represent complicated functions in simple terms.

Infinite sums

Euler introduced a special sign, which continues to be used today, to signify a sum total or 'summation'. It was Σ, the capital of the letter *sigma* (s) in the Greek alphabet and also the first letter in the word *sum*.

A summation is an expression of the type:

$$\sum_{i=1}^{i=5} i$$

that specifies a variable, in this case i, with subscripts and superscripts telling us how the variable in question changes. In the example, these subscripts and superscripts tell us that i varies from 1 to 5. That is:

$$\sum_{i=1}^{i=5} i = 1+2+3+4+5;$$

$$\sum_{n=1}^{n=3} (n+1) = (1+1)+(2+1)+(3+1);$$

$$\sum_{n=1}^{n=4} n^2 = 1^2+2^2+3^2+4^2.$$

It is common to simplify the notation by placing only the last term in the series above the sigma, thus:

$$\sum_{i=1}^{5} i = 1+2+3+4+5$$

indicating that i varies from 1 to 5.

If the upper limit is not given a number, then the infinity symbol is used, showing that the sum is infinite. For example:

$$\sum_{i=1}^{\infty} i = 1+2+3+4+5+\ldots$$

Although it may initially seem strange, finite sums exist, the final result of which is a finite number. Such a series would be described as 'convergent'. For example, the series:

$$\sum_{i=1}^{\infty} \frac{i}{2^i} = \frac{1}{2}+\frac{2}{4}+\frac{3}{8}+\frac{4}{16}+\ldots$$

has a finite sum with a value of nearly 2. As the terms get smaller and smaller, there will come a point where one is so close to zero that its addition will hardly change the total, and the final sum will be a finite number. This is one way of looking at it, but it lacks precision. We might suppose that a series of the type:

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$$

also has a finite sum, but this is not so. This particular series, in which Euler was especially interested, is called a 'harmonic'. He used it to obtain a proof different from the one he had already deduced to prove the existence of an infinity of prime numbers.

THE BASEL PROBLEM

Jacques Bernoulli (1654–1705) and his brother Johann (1667–1748) devoted themselves to the study of harmonic series, especially between the years 1689 and 1704. It was they who proved that some series are divergent. Encouraged by the results, they studied the series given by inverse squares:

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Jacob showed that the series converged and even managed to prove that the sum should be less than or equal to 2, but he failed to find the exact value of this series. He was so taken with this problem that he said, "Great will be our gratitude if anyone finds and informs us of what has hitherto escaped our endeavours." The matter was known as the 'Basel problem', as it was at the university in this Swiss city that Johann held the chair in mathematics and from where he made his famous proposal.

Many great mathematicians including Mengoli and Leibniz failed to conquer the challenge, not to mention the joint efforts made by the Bernoulli brothers. The solution, which arrived 30 years later, was found by Euler, the 'Magician'. The result was truly spectacular:

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots = \frac{\pi^2}{6}.$$

A harmonic series diverges, meaning that the sum of its terms is infinite, but our example does so extraordinarily slowly in comparison with a series of the type:

$$\sum_{n=1}^{\infty} n^2 = 1^2 + 2^2 + 3^2 + 4^2 + \ldots$$

Working with the harmonic series, Euler devised a function that would go down in history as one of the most important to be established in mathematics, 'Euler's zeta function' – although nowadays it is also known somewhat unfairly as the 'Riemann zeta function'. To describe it, Euler used the Greek letter ζ (zeta):

$$\zeta(x) = \frac{1}{1^x} + \frac{1}{2^x} + \frac{1}{3^x} + \frac{1}{4^x} + \ldots + \frac{1}{n^x} + \ldots$$

Euler wrote the following about it:

"... However, I have now discovered, against all expectations, an elegant expression for the sum of the series 1 + 1/4 + 1/9 + 1/16 + ..., which depends on squaring the circle ... I have found that six times the sum of this series is equal to the square of the length of the circumference of a circle whose diameter is one."

Unfortunately, Jacob had died by the time Euler published his result. "If only my brother had lived!" Johann is said to have lamented.

The title 'Magician' given to Euler is due to the seemingly magical maths that the proof is supposed to demonstrate. In fact, it is not at all complicated but does require some knowledge of higher mathematics, and reflects Euler's audacity to treat the series in question as if it were a polynomial function and then relate it to the progression of the sine function; this accounts for the appearance of π, which is one of the zero values of this function.

Johann Bernoulli was Euler's teacher and one of the best mathematicians of his day.

If we set $x = 1$ we obtain the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$, which we have previously encountered, and the sum of its terms we know is infinite. However, Euler suspected that, if he set $x = 2$, the resulting series

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots,$$

would not be infinite, as he had taken only the fractions containing the squares from the harmonic series. To calculate the value of the latter series was practically impossible with the knowledge of that time. However, in one of his most brilliant discoveries, Euler succeeded in proving the following equality:

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots = \frac{\pi^2}{6}$$

Euler made this discovery when he was 28 years old, although it took him a further six years to perfect the proof. The sudden appearance of π, used in measuring the dimensions of a circle, in the result of this sum caused astonishment throughout the mathematics community of the period. With this finding Euler solved one of the most intriguing problems of the time, the so-called 'Basel problem'.

Experimenting with the zeta function, Euler obtained a number of results. What he knew for certain was that, when x has a value less than or equal to 1, the value of the sum is infinite, and that therefore the series only converges for values of x that are greater than 1.

EULER AND SOUND

Euler decided to introduce an imaginary variable into the so-called exponential function $f(x) = 2^x$. He was amazed to find that the graph of this function contained waves, a series of undulating lines that were the same as those encountered in attempts to represent musical notes. According to the values taken by these imaginary numbers, the waves corresponded to notes of higher or lower pitch.

Several years later, Jean-Baptiste-Joseph Fourier (1768–1830), a mathematician of French origin, exerted a strong influence on mathematical physics through his system known as the Fourier series. It was based on Euler's result for analysing periodic functions, which closely relates analytical methods to the world of sound.

Euler then thought of associating prime numbers with the function. He knew that Euclid's fundamental theorem of arithmetic stated that any natural number can be expressed as a product of prime numbers in one way only. This meant that each of the fractions involved in the zeta function could be written so that the denominator contained only prime numbers. For example, suppose that we give the zeta function the value $x = 2$:

$$\zeta(2) = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots + \frac{1}{n^2} + \ldots$$

and we take n = 360 in order to calculate the inverse of 360^2.
We reduce the number 360 to prime factors: $360 = 2^3 \cdot 3^2 \cdot 5$, so that:

$$\frac{1}{360} = \frac{1}{2^3} \cdot \frac{1}{3^2} \cdot \frac{1}{5^1}.$$

Squaring all the terms, we obtain:

$$\left(\frac{1}{360}\right)^2 = \left(\frac{1}{2^3}\right)^2 \cdot \left(\frac{1}{3^2}\right)^2 \cdot \left(\frac{1}{5^1}\right)^2.$$

By doing this with each of the denominators of the zeta function, Euler obtained the expression:

$$\zeta(x) = \frac{1}{1^x} + \frac{1}{2^x} + \frac{1}{3^x} + \frac{1}{4^x} + \ldots + \frac{1}{n^x} + \ldots =$$

$$= \left(1 + \frac{1}{2^x} + \frac{1}{4^x} + \frac{1}{8^x} \ldots\right) \cdot \left(1 + \frac{1}{3^x} + \frac{1}{9^x} + \frac{1}{27^x} \ldots\right) \cdot \left(1 + \frac{1}{p^x} + \frac{1}{(p^2)^x} + \frac{1}{(p^3)^x} \ldots\right) \ldots,$$

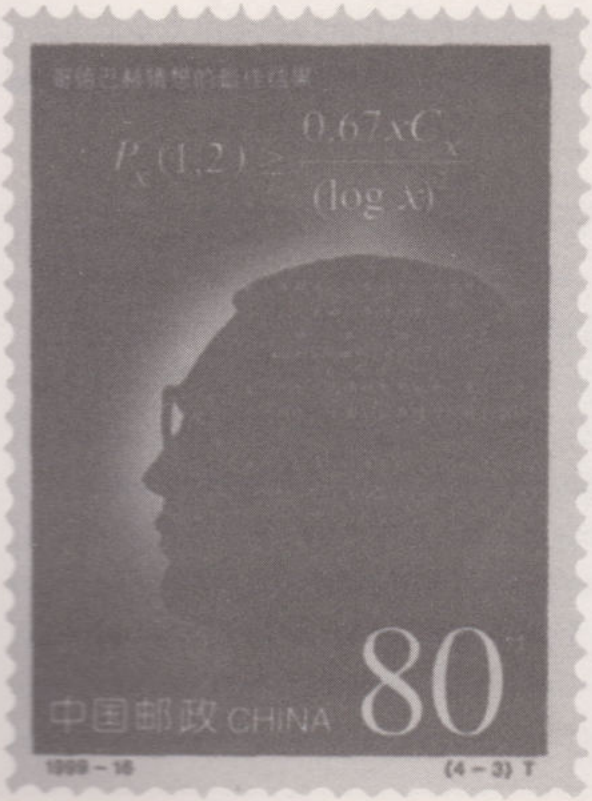
which contains only prime numbers. This equation has a left-hand term that is an infinite sum and a right-hand term that is a product, also of an infinite set of numbers. This expression can be considered the foundation stone of what would become the edifice of analytical number theory developed in later centuries. This expression, known as the 'Euler product', was the starting point used by Riemann to begin to impose order on the chaotic cohort of prime numbers, as we shall see in Chapter 6.

The Goldbach conjecture

Christian Goldbach (1690-1764) was a Prussian mathematician who frequently corresponded with Euler. On 18 November 1752, he sent him a letter containing the following proposition: "Any even number greater than 2 can be written as a sum of two prime numbers." The expression 'sum of two primes' used here includes cases in which a prime number is repeated. For example,:

$$4 = 2 + 2$$
$$6 = 3 + 3$$
$$8 = 3 + 5$$
$$10 = 3 + 7$$
$$12 = 5 + 7$$
$$14 = 3 + 11.$$

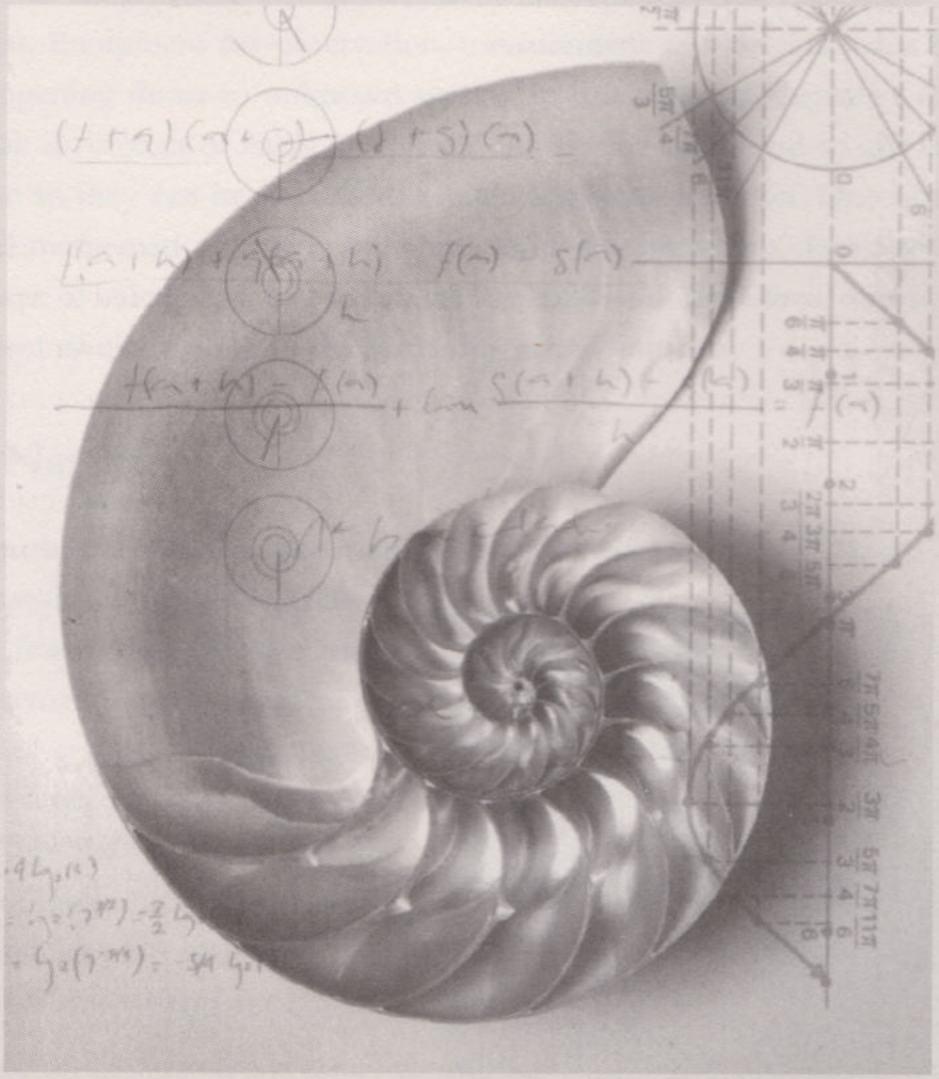
On 16 December that year, Euler replied saying that he had verified the conjecture up to the number 1,000, and in another letter dated 3 April 1753 that he had checked the result as far as the number 2,500. At present, the conjecture has been verified by computers for all even numbers up to two trillion. However, the conjecture has not yet been proved and it is considered by the mathematics community to be one of the most intractable problems in the history of science.



Chen Jingrun (1933–1996), one of the most outstanding mathematicians of the 20th century, provided the best result yet of the Goldbach conjecture in 1966 when he showed that any even number that is sufficiently large can be written as the sum of a prime and a semi-prime (a number that is the product of just two prime factors). This fact is demonstrated on the postage stamp with which the People's Republic of China commemorated Chen in 1999. His inequality appears above his silhouette.

UNCLE PETROS AND GOLDBACH'S CONJECTURE

This is the title of a famous novel by Apostolos Doxiadis in which a retired mathematician asks his nephew to solve a mathematics problem. The protagonist's intention is that his nephew should give up studying mathematics at university if he fails to solve the problem during his vacation. Having spent the entire summer trying to solve it, the nephew gives up and graduates in law. The problem set was Goldbach's conjecture. In an effort to generate publicity for the novel, in 2000 the British publisher Tony Faber offered a prize of one million dollars to any English-speaking person, who could prove the conjecture by April 2002. As expected, nobody claimed the prize.



An illustration of the cover of some editions of Apostolos Doxiadis's book, featuring a nautilus shell, a naturally occurring logarithmic spiral.

Chapter 4

Logarithms and Prime Numbers

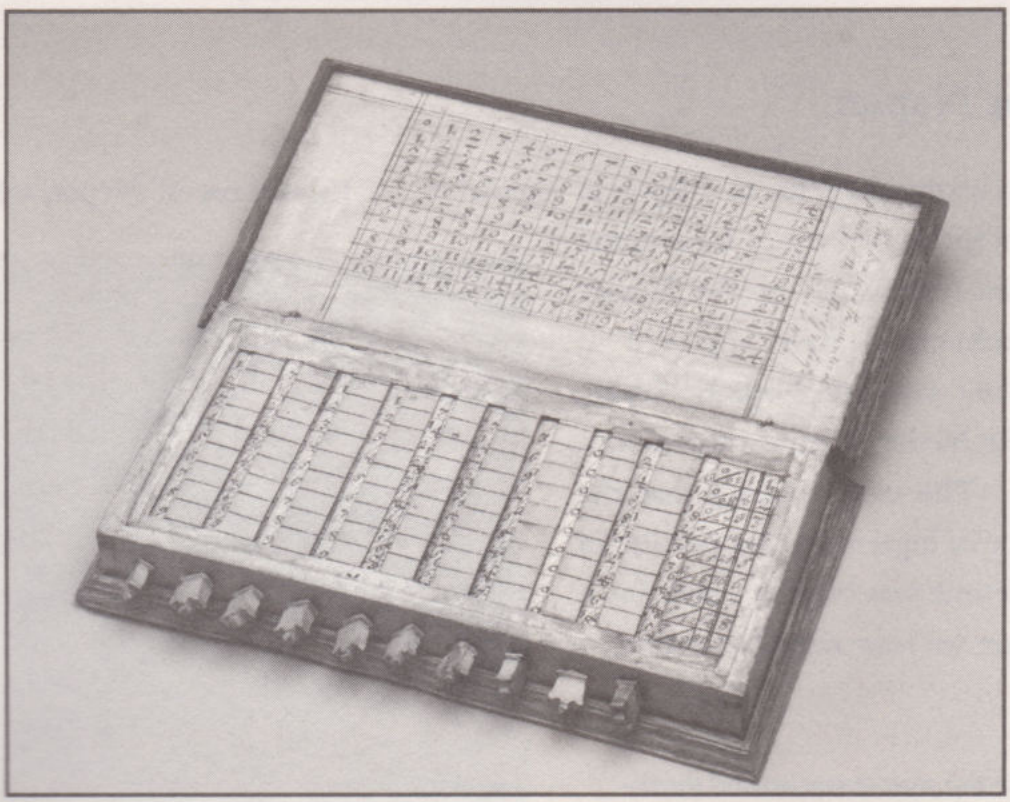
When we investigate an object, the devices we use to observe it have an impact on what we can see. For example, the development of astronomy was closely linked with the development of telescopes, just as microbiology was associated with microscopes. Equipment for observation, measurement and detection has been the key to opening doors to unknown worlds. In this sense, mathematics is not exceptional: its objects of study are found only in the mind and so are intangible but, even so they can be defined to a high degree of precision. One of the most powerful mathematical devices ever invented is the logarithm, developed initially as a means of calculation but which, in the hands of Carl Gauss, ended up as an instrument of observation in the search for prime numbers.

John Napier

Many textbooks refer to Neperian logarithms or logarithms of Neper, whereas others speak of Napier logarithms. In fact, few names in the history of mathematics have appeared in so many variants: Napeir, Nepair, Nepeir, Neper, Napare, Naper, Naipper,... The only spelling that the creator of logarithms did not use in his lifetime was Napier – and this is the one we now regard as correct!

The Scottish mathematician and theologian, John Napier, has gone down in history for his methods of simplifying complex calculations.

John Napier was born in 1550 at Merchiston Castle near Edinburgh in Scotland. He was the son of an aristocrat, Archibald Napier, who lived in very comfortable circumstances. John studied theology at the University of St. Andrews. His interest in mathematics can be traced to a long journey he made around Europe. It is recorded that he was at university in Paris and also spent some time in Italy and Holland. On returning to Scotland in 1572 he married Elisabeth Stirling. He then devoted the next two years to building a castle in Gartness. Napier spent a lot of time shut up in this castle and it was during this period that he became engrossed in his mysterious mathematical activities. We say 'mysterious', because on the few occasions that Napier appeared in public he was dressed all in black and carried a cockerel, also black, on his shoulder. This eccentricity earned him the reputation of being a wizard, something that was enhanced by his demonstration of mathematical skills that no one else possessed. In addition to being exceptionally fond of mathematics, he spent much of his time examining the gospels, especially the *Book of Revelation*. He published his conclusions in a work called *Plaine Discovery of the Whole Revelation of Saint John*, which was translated into a number of languages, and in which the author attempted to demonstrate that the pope was the antichrist.



One of the first models of the Naperian abacus, also known as Napier's Bones, invented by John Napier to calculate products and quotients.

STRANGE DECIMALS

Being able to express a fraction such as 19/8 as a decimal 2.375 seems perfectly normal to us – we simply divide 19 by 8. But in the 16th century decimal expressions were still very exotic. The Flemish engineer Simon Stevin (1548–1620) introduced a decimal fraction notation and suggested weight and measurement units that were based on decimals, like the metric system largely in use today. Napier supported the use of decimal fractions and simplified Stevin's notation, introducing a comma as a separator (technically known as the 'radix point') between the integers and the decimal fractions. The comma is still used in many European countries. However, in English-speaking countries, a dot is used as the decimal separator.

Napier was interested in numerology and astrology. The second interest led him to investigate the properties of geometrical figures on a spherical surface, and he obtained important results for solving spherical triangles. Any student who has tackled spherical trigonometry will have encountered more than one formula bearing the Scotman's name.

However, for Napier there was one issue that would end up overriding all others. In those days, numerical calculations were extremely tedious. Napier thought that he could use his time more profitably than simply filling up sheet after sheet of paper with endless calculations that were actually just routine work.

He succeeded in inventing a device, based on rods with a square cross section that fitted into multiplication boards and made it possible to carry out additions and multiplication quite easily. In 1617 he published a handbook entitled *Rabdologiae* (*The Study of Rods*) in which he explained how the device should be used. Napier's tool, the precursor of the slide rule, was used in Scotland for more than 100 years. (Napier later improved the device by replacing the rods with punch cards that enabled the multiplication of much larger numbers. In fact, it could be said that these cards are a clear antecedent – among others – of the famous punch cards that would appear more than four centuries later with the first IBM computers.)

However, Napier's greatest creation in terms of the history of mathematics was logarithms, an ingenious method of calculation that he published in 1614 under the title *Mirifici Logarithmorum Canonis Descriptio* (*Construction of the Marvellous Canon of Logarithms*). To appreciate the central role that logarithms have achieved in the study of prime numbers, we will first look at some of their basic concepts.

Logarithms

Logarithms are based on the following idea: We know that $1{,}000 = 10 \cdot 10 \cdot 10$, that is, ten raised to the power three, which can be expressed as 10^3, so that:

$$1{,}000 = 10^3$$
$$10{,}000 = 10^4$$
$$1{,}000{,}000 = 10^6.$$

Suppose that we wish to multiply these numbers together:

$$1{,}000 \cdot 10{,}000 \cdot 1{,}000{,}000 = 10{,}000{,}000{,}000{,}000$$
$$\text{But } 10{,}000{,}000{,}000{,}000 = 10^{13}$$

We could have performed the multiplication by writing $10^{3+4+6} = 10^{13}$. It is obviously easier to add than to multiply. Just to be sure, try multiplying $10^{38} \cdot 10^{52} = 10^{90}$ by writing out the numbers in full!

We now move on to the language of logarithms. Using the example of $1{,}000 = 10^3$, we may ask 'to what power do we raise the number 10 to arrive at 1,000?' The answer is 3. We write this in the following way: $\log (1{,}000) = 3$.

Hence, for example:

$$\log 100 = 2$$
$$\log 1{,}000 = 3$$
$$\log 1{,}000{,}000 = 6.$$

The concept underlying this pattern is that it is much easier to add than to multiply. For example:

$$\log (100 \cdot 1{,}000) = \log 100 + \log 1{,}000 = 2 + 3 = 5.$$

Then we simply proceed in reverse, using the antilogarithm, to obtain the final result: $10^5 = 100{,}000$.

We could demonstrate all these operations in a table, like the one below:

1	10	100	1,000	10,000	100,000	1,000,000	10,000,000	100,000,000	1,000,000,000
0	1	2	3	4	5	6	7	8	9

The top row of the table starts with the number 1 and each successive box is equal to ten times the previous one. This is called a geometric progression by a factor of 10. In contrast, the figures in the bottom row are obtained by adding 1 to the number in the preceding box. Thus, the top row contains products and the bottom row contains additions. According to the table:

$$1{,}000 \cdot 100{,}000 = 100{,}000{,}000$$

and is equivalent to the sum:

$$3 + 5 = 8.$$

We can write such a table by entering any geometric progression in the top row, for example:

1	2	4	8	16	32	64	128	256
0	1	2	3	4	5	6	7	8

To multiply 4 by 16 (top row) we add 2 and 4 (bottom row) to give 6, which corresponds to the answer 64. Similarly, we can perform divisions, but in that case the result is obtained by subtracting the corresponding numbers in the bottom row. For example, to divide 256 by 8, we simply subtract 3 from 8, i.e. $8 - 3 = 5$, giving 32, which is the number in the box above 5. This relationship between the numbers in the bottom and top rows is the key to logarithms.

We can now posit a rigorous definition of a logarithm. When we say that the number 32 corresponds to 5, what we are expressing is the equality:

$$2^5 = 32.$$

Remember that 2 raised to the power 5 means that 2 is multiplied by itself five times. We can read the two rows of the second table as follows: '3 is the number to which 2 has to be raised to give the answer 8' and '7 is the number to which 2 has to be raised to give the answer 128', this can be abbreviated to:

$$\log_2 8 = 3$$
$$\log_2 128 = 7$$

These expressions are read as 'the logarithm to base 2 of 8 is 3' and 'the logarithm to base 2 of 128 is 7', respectively. Now take an example from the first table, $10^4 = 10{,}000$, i.e. 4 is the power to which 10 has to be raised to get 10,000. Expressing this as a logarithm we get $\log_{10} 10{,}000 = 4$, which is read as 'the logarithm to base 10 of 10,000 is 4'.

This means that we can establish the general definition of a logarithm. The logarithm to base a of a number b is the number c to which the base a has to be raised to yield b, ($a^c = b$), and we write this as:

$$\log_a b = c.$$

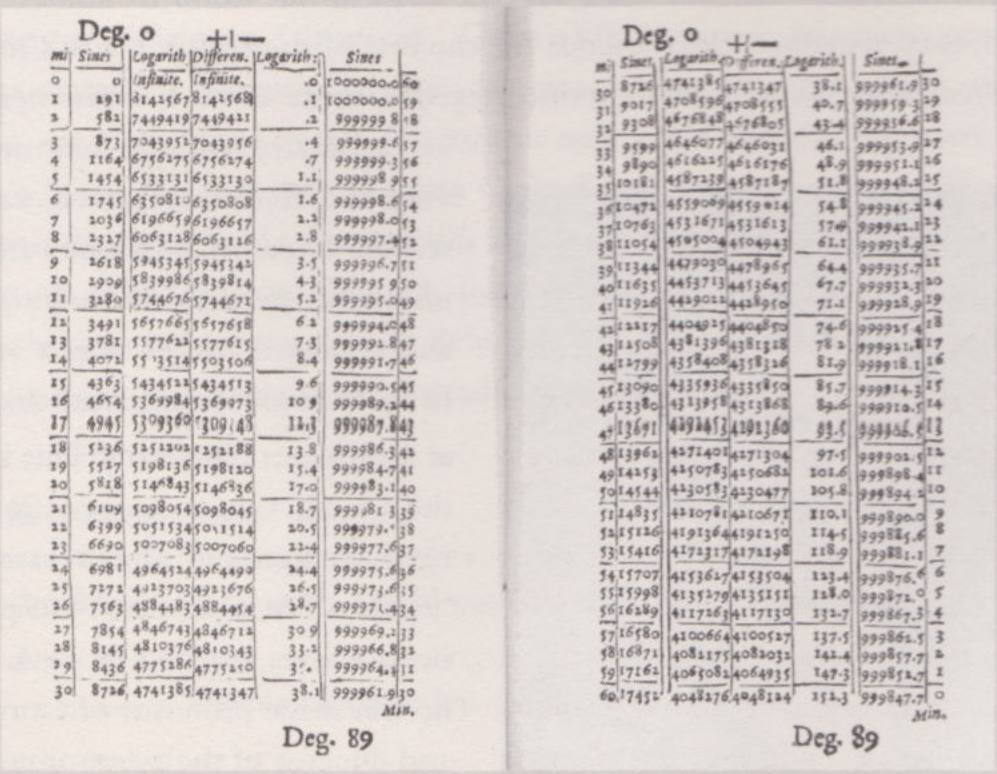
Napier was interested in simplifying calculations in spherical trigonometry and his notion of a logarithm was first applied to trigonometric functions. His original approach was not like ours, which could be called arithmetical, but was 'kinematic', whereby he posited two straight-line segments traversed at different speeds. The word 'logarithm' was first used by Napier himself and means 'number of ratio' in reference to the relationship between the different straight-line segments used by him. (In our case, it is the relationship between the numbers in the two rows of the tables). Napier worked with logarithms to base 10^7, which was not enormously practical. Furthermore, he failed to establish that the logarithm of 1 is 0, which is tantamount to saying that $10^0 = 1$. Henry Briggs (1561–1630), was Professor of Geometry at Oxford University, wrote to him expressing the interest he had developed in logarithms and suggested that they meet. In the summer of 1615 Briggs went to see Napier at Merchiston Castle, and they discussed the possibility of using the number 10 as a base and of log $1 = 0$. Napier, who was by then a sick man, refused to draw up a new version of his logarithmic tables. Napier died the next year and Briggs then devised a definition of a logarithm very similar to that given here, thus establishing what are known as 'Briggs logarithms'.

However, an apparently fortuitous incident in the preparation of logarithmic tables would mark a milestone in the history of mathematics. Just as in school books it was customary to include multiplication tables on the back cover, so a list of prime numbers was placed at the end of most logarithmic tables. There was a good reason for this. If we recall that any number can be expressed as a product of prime factors, it is logical to calculate the logarithm of prime numbers first and then obtain the logarithms of other numbers by simple addition.

The logarithmic tables used by Gauss at school contained a list of the first 1,000 prime numbers. A brilliant mind was confronted by two apparently unconnected concepts and their subsequent combination resulted in one of the most interesting theorems in algebra.

LOGARITHMIC TABLES

Nowadays, calculating a logarithm is as simple as pressing a key on a pocket calculator, but in the 17th century large books containing the logarithms of as many numbers as possible were used. In 1617 Henry Briggs published the first tables to contain the logarithms of the numbers between 1 and 1,000 to an accuracy of 14 decimal places. Seven years later new tables appeared, first with values between 1 and 20,000 and then between 90,000 and 100,000, also to 14 decimal places. Editions of these tables were soon being printed in other countries in view of the enormous practical value of calculating with logarithms. Maritime navigation required ever more accurate astronomical maps, and the complexity of the trigonometric calculations these demanded meant that astronomers would spend hours, days and even years on the task. As Pierre-Simon Laplace put it: "Thanks to his [Napier's] efforts, the lifetime of astronomers has been doubled".


Deg. o

mi	Sines	Logarith	Differen.	Logarith	Sines	
		Infinite.	Infinite.	.0	1000000.0	60
1	291	8142567	8142568	.1	1000000.0	59
2	582	7449419	7449421	.2	999999.8	58
3	873	7043951	7043956	.4	999999.6	57
4	1164	6756275	6756274	.7	999999.3	56
5	1454	6533131	6533130	1.1	999998.9	55
6	1745	6350811	6350808	1.6	999998.6	54
7	2036	6196659	6196657	2.2	999998.0	53
8	2327	6063128	6063126	2.8	999997.4	52
9	2618	5945345	5945342	3.5	999996.7	51
10	2909	5839986	5839814	4.3	999995.9	50
11	3280	5744676	5744672	5.2	999995.0	49
12	3491	5657665	5657658	6.2	999994.0	48
13	3781	5577622	5577615	7.3	999992.8	47
14	4072	55 3514	5503506	8.4	999991.7	46
15	4363	5434522	5434513	9.6	999990.5	45
16	4654	5369984	5369973	10.9	999989.2	44
17	4945	5309360	5309343	12.3	999987.8	43
18	5236	5252202	5252188	13.8	999986.3	42
19	5527	5198136	5198120	15.4	999984.7	41
20	5818	5146843	5146836	17.0	999983.1	40
21	6109	5098054	5098045	18.7	999981.3	39
22	6399	5051534	5051514	20.5	999979.5	38
23	6690	5007083	5007060	22.4	999977.6	37
24	6981	4964124	4964499	24.4	999975.6	36
25	7272	4923703	4923676	26.5	999973.6	35
26	7563	4884183	4884454	28.7	999971.4	34
27	7854	4846743	4846712	30 9	999969.2	33
28	8145	4810376	4810343	33.2	999966.8	32
29	8436	4775286	4775250	35.6	999964.4	31
30	8726	4741385	4741347	38.1	999961.9	30

Min.

Deg. 89

Deg. o

m.	Sines	Logarith	Differen.	Logarith	Sines	
30	8726	4741385	4741347	38.1	999961.9	30
31	9017	4708596	4708555	40.7	999959.3	29
32	9308	4676848	4676805	43.4	999956.6	28
33	9599	4646077	4646031	46.1	999953.9	27
34	9890	4616227	4616176	48.9	999951.1	26
35	10181	4587238	4587187	51.8	999948.2	25
36	10472	4559069	4559014	54.8	999945.2	24
37	10763	4531671	4531613	57.9	999942.1	23
38	11054	4505004	4504943	61.1	999938.9	22
39	11344	4479030	4478965	64.4	999935.7	21
40	11635	4453713	4453645	67.7	999932.3	20
41	11926	4429021	4428950	71.1	999928.9	19
42	12217	4404925	4404850	74.6	999925.4	18
43	12508	4381396	4381318	78.2	999921.8	17
44	12799	4358408	4358326	81.9	999918.1	16
45	13090	4335936	4335850	85.7	999914.3	15
46	13380	4313958	4313868	89.6	999910.5	14
47	13671	4292453	4292360	93.5	999906.5	13
48	13962	4271401	4271304	97.5	999902.5	12
49	14253	4250783	4250682	101.6	999898.4	11
50	14544	4230583	4230477	105.8	999894.2	10
51	14835	4210781	4210671	110.1	999890.0	9
52	15126	4191364	4191250	114.5	999885.6	8
53	15416	4172317	4172198	118.9	999881.1	7
54	15707	4153617	4153504	123.4	999876.6	6
55	15998	4135279	4135151	128.0	999872.0	5
56	16289	4117263	4117130	132.7	999867.3	4
57	16580	4100664	4100527	137.5	999862.5	3
58	16871	4082175	4082032	142.4	999857.7	2
59	17162	4065082	4064935	147.3	999852.7	1
60	17452	4048276	4048124	152.3	999847.7	0

Min.

Deg. 89

The first logarithmic tables were published in Edinburgh in 1614.

Johann Carl Friedrich Gauss

Gauss was born in Braunschweig, Germany, on 30 April 1777. His background was a poor one and he was destined for farm work, unless fate intervened. However, at primary school Gauss, aged just nine, was already an outstanding pupil. His was a country school with few resources where the single teacher, one Herr Büttner, had to keep around 100 pupils occupied. One way of doing this was to set them tedious routine calculations. On one occasion he made them work out the sum of the first 100 natural numbers. In a moment Gauss put down his exercise book and said "Here you are!" Gauss had not only carried out the addition

$$1 + 2 + 3 + 4 + \ldots + 100 = (1 + 100) + (2 + 99) + (3 + 98) + \ldots + (50 + 51)$$
$$= 101 + 101 + \ldots + 101 = 101 \cdot 50 = 5{,}050$$

in record time but he had also solved the problem of adding the terms as an arithmetical progression. Büttner immediately saw that this was an exceptionally gifted pupil and decided to pass him on to Johann Martin Bartels (1769–1836), a keen student of mathematics, himself just eight years older than Gauss. With Bartels the boy took his first steps in the world of numbers and, indeed, they remained close friends for the rest of their lives. Gauss's mother, Dorothea Benz, aware that something needed to be done to assist her son's extraordinary abilities but not having the resources herself, contacted the Duke of Braunschweig. The nobleman became the boy's patron and managed to get him a stipend to study at grammar school and later at Göttingen University. This is how the young Gauss escaped from his rural background and became the 'Prince of Mathematics'. His professional career reached its peak when he was made professor of astronomy and director of the astronomical ob-

Portrait of Johann Gauss as a young man.

A famous lithograph by Eduard Ritmüller depicts the by now distinguished Gauss on the terrace of the observatory at Göttingen University.

servatory at Göttingen University. Gauss's life was spent relatively peacefully. He maintained a conservative attitude towards the duke, his patron, in what was an age of political unrest. He was an only child and did not marry until the age of 32. His wife was Johanna Osthoff and they had three children, the youngest of whom died a few months after Johanna's own death. In 1810 Gauss remarried, to Wilhelmine Waldeck, the daughter of a law professor, and three more children were born. On 23 February 1855 Gauss died in Göttingen. By then, his fame as a scientist had spread around the world.

The first conjecture

A notebook used by Gauss at age 14 contains the remark:

"Prime numbers less than a $(= \infty)$ a/la."

Gauss had focused on studying the long list of prime numbers appearing at the end of his logarithmic tables, and had become captivated by the chaotic nature of

the series. However, he had already decided that it was not his vocation to find a formula indicating how and where the next prime number would occur. He clearly felt that this task would end in failure. Instead, he proceeded to calculate how many prime numbers there are between two given numbers or, more precisely, how many prime numbers there are among the first 10, 100, 1,000 and 10,000 numbers, as this would allow him to estimate the frequency with which prime numbers appear in the sequence of natural numbers.

We know that the first 10 natural numbers contain only four prime numbers (2, 3, 5 and 7). Between 10 and 100 there are 21 more. To express this, Gauss defined a function that he called $\pi(x)$ and expressed as follows:

$$\pi(x) = \text{the number of primes that are smaller than } x.$$

A SCIENTIST THROUGH AND THROUGH

Gauss also worked in several areas outside mathematics. Noteworthy are the results he obtained in work on the Earth's magnetic field, electromagnetism, capillarity, the attraction of ellipsoids and dioptrics. In the field of geodesy, Gauss invented the heliostat (a device that could send signals – and precise locations – using reflected light). A curious event involving his research occurred in 1833, when Gauss was working with Wilhelm Weber (1804–1891) on electromagnetism. To be able to send messages quickly, Gauss personally built an electrical device capable of transmitting messages at the speed of light. He had invented nothing less than the electric telegraph.

Monument to Gauss and Weber in Göttingen.

According to this, $\pi(10) = 4$.

For example, to calculate $\pi(15)$ we need to count the prime numbers smaller than 15, i.e.:

$$2, 3, 5, 7, 11, 13.$$

and so $\pi(15) = 6$.

The symbol π appearing in the formula is more famous as the constant pi (3.14159...), but in this context it lacks that mathematical significance. The function could be defined just as well by any other symbol, for example $C(x)$. Indeed, the choice of π by the young Gauss was not a good one, and it is likely that he just wrote down the first symbol that came into his head. To most of us, the term $\pi(x)$ automatically suggests some kind of mathematical connection with the circumference of a circle but this has nothing to do with prime numbers in the present context. In any event, we will continue to use Gauss's notation here.

The German mathematician then constructed a two-column table, putting powers of 10 in the left-hand column and the values taken by $\pi(x)$ on the right.

The following table has been calculated for the first 10 billion numbers. Clearly, in Gauss's day calculation methods were much less reliable and he did not have a similar range of values:

x	$\pi(x)$
10	4
100	25
1,000	168
10,000	1,229
100,000	9,592
1,000,000	78,498
10,000,000	664,579
100,000,000	5,761,455
1,000,000,000	50,847,534
10,000,000,000	455,052,512

Logically, $\pi(x)$ is a number that will go on increasing, but the way in which it does so will not tell us much. We will add another column to give us the proportion of the numbers smaller than a given number that are primes. To do this we calculate the quotient:

$$\frac{\pi(x)}{x}.$$

We know that there are 168 prime numbers smaller than 1,000:

$$\frac{\pi(x)}{x} = \frac{\pi(1{,}000)}{1{,}000} = \frac{168}{1{,}000} = 0.168.$$

This number informs us that 16.8% of numbers between 1 and 1,000 are primes. The remaining 83.2% consist of composite numbers. When we add this third column to the table

x	$\pi(x)$	$\pi(x) / x$
10	4	0.40000000
100	25	0.25000000
1,000	168	0.16800000
10,000	1,229	0.12290000
100,000	9,592	0.09592000
1,000,000	78,498	0.07849800
10,000,000	664,579	0.06645790
100,000,000	5,761,455	0.05761455
1,000,000,000	50,847,534	0.05084753
10,000,000,000	455,052,512	0.04550525

we see that the proportion of prime numbers decreases as we progress to ever larger numbers. This is an important, indeed a predictable, fact. A number is only prime if it is not divisible by any of the numbers preceding it, except for 1. For example, for 13 to be prime it must not be divisible by 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12. The larger the number, the greater the number of possible divisors, and hence it is logical that prime numbers will become scarcer. But Gauss already knew this

did not mean that all prime numbers would eventually cease, as he was perfectly aware of the existence of the fundamental theorem of arithmetic in which Euclid had demonstrated that the set of prime numbers is infinite.

The third column that Gauss included in the table was not obtained from the quotient $\frac{\pi(x)}{x}$ but from its inverse, $\frac{x}{\pi(x)}$.

x	$\pi(x)$	$x / \pi(x)$
10	4	2.5
100	25	4
1,000	168	6
10,000	1,229	8.1
100,000	9,592	10.4
1,000,000	78,498	12.7
10,000,000	664,579	15
100,000,000	5,761,455	17.4
1,000,000,000	50,847,534	19.7
10,000,000,000	455,052,512	22

This table tells us, for example, that in the first 100 numbers, one in four is prime, that in the first 1,000 numbers about one in every six is prime, and so on. This, of course, is an estimate. The table does not assert that in the first 100 numbers every fourth number is prime, a fact that can be quickly checked by applying Eratosthenes' sieve to the first 100 numbers. Therefore, the above table should only be seen as indicating the probable distance between prime numbers.

Gauss noted that the last column grew by approximately two units in each successive row. The situation then is as follows: as each row is increased by a factor of 10, two more units are added to the proportion of its primes. This relationship between product and sum is inherent in the very nature of logarithms. Gauss had a table of logarithms and another of prime numbers in the same book. This gave him the idea of devising a new observational tool. Logarithms would become another lens to be fitted to his mathematical telescope. As we have already seen, when logarithms are to base 10, every time we multiply by 10, decimal logarithms increase by 1, meaning that this base does not fit well with Gauss's scheme, and so

he decided to take logarithms to the base e, a number with similar features to π. Its approximate value is:

$$e = 2.71828182845904523 54...$$

This is an infinite decimal number that appears in mathematics about as often as π does. When a logarithm is taken to the base e, we say we are dealing with 'natural logarithms'. For reasons we have explained, we could write $\log_e$ to indicate natural logarithms; however, scientific calculators have a key for *log*, referring to decimal logarithms, and another for *ln* (log to the base *e*).

The conjecture raised by Gauss on this basis is: for large values of x, the value of $\frac{\pi(x)}{x}$ approximates to $\frac{1}{\ln x}$, which can also be expressed as:

$$\frac{\pi(x)}{x} \approx \frac{1}{\ln x} \text{ (for values greater than } x).$$

This result is an estimate of the frequency with which prime numbers appear in the sequence of natural numbers. Suppose that $P(N)$ is the number of primes smaller than N. The formula states that, as N increases, the quotient $N/P(N)$ gets closer and closer to the natural logarithm of N.

There is a simple way of implementing Gauss's formula if we want to know how many prime numbers there are that are smaller than a given number. For example, suppose that someone asks us the following question: 'How many prime numbers do you think there are in the first 1,000 natural numbers?'

We take a pocket calculator and proceed as follows:

1) Enter the number 1,000;
2) Press the ln key;
3) Press the $1/x$ key;
4) Multiply the result by 1,000, and the number
144.764827301083942550376306305 54

appears, enabling us to give an answer to the question: 'There are about 145 prime numbers between 1 and 1,000'

This, of course, is only an approximation as the real answer is 168. However, we should bear in mind that the theorem becomes more accurate as the number N increases, and we can say with some confidence, for example, that only 3.6% of the first billion natural numbers are prime.

We can now decipher what Gauss meant when he wrote "prime numbers less than a $(= \infty)$ a/la" in his notebook:

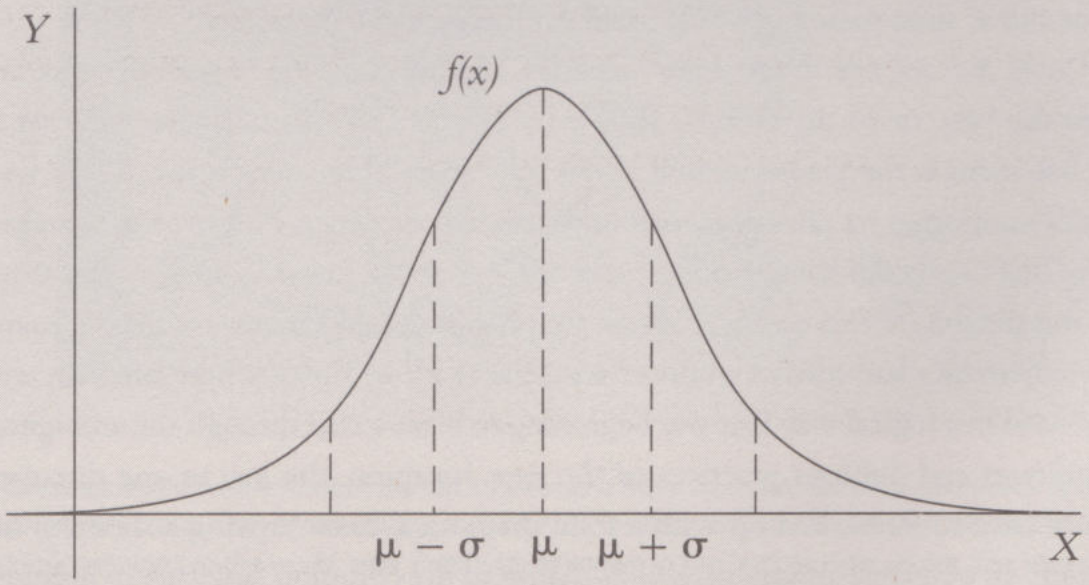
'prime numbers less than a' is the same as saying '$\pi(a)$';

'la' is what we now write as $\ln a$;

'$= \infty$' means that the equality is most valid for very large values of a (when a tends to infinity).

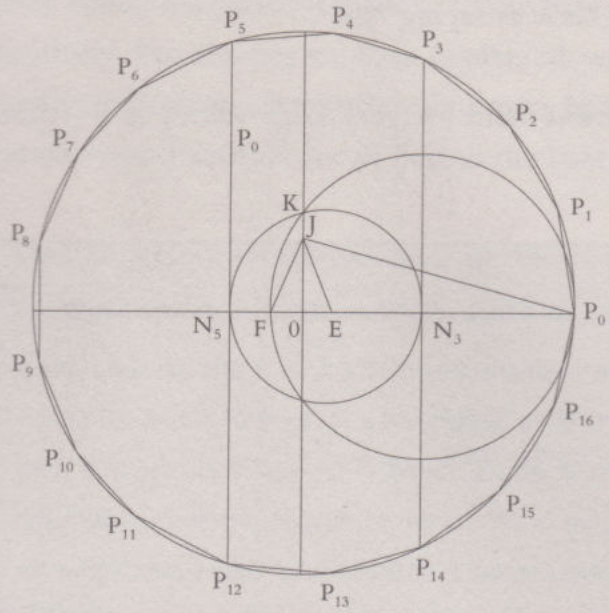
THE GAUSSIAN BELL CURVE

When he was 18, Gauss discovered the least-squares method, and this caused him to take a special interest in error theory. He then developed a method of statistical observation in which the normal distribution of errors traced out a bell-shaped curve. This is without a doubt the best-known curve in mathematics, and is generally called the 'Gaussian bell curve'. This method of observation eventually paid Gauss handsome returns, as he began a systematic study of international stock market trends appearing in foreign newspapers that regularly turned up in the university's lecture theatres. Gauss's bell curve was a success, and the profits he made as a result of this research greatly exceeded his professor's salary.



GAUSS'S POLYGON

The construction of regular polygons with a ruler and compass had been an unresolved problem since the time of the Greek geometers. Shapes with 3, 4, 5 and 15 sides, and double these, could be constructed. On 30 March 1796, exactly a month before his 20th birthday, Gauss found how to construct a polygon with 17 sides using only a ruler and compass.



It was an important date in his career as, that very same day, he started a scientific diary covering the years 1796–1814, which is considered to be a veritable jewel in mathematics, as it contains notes on all his scientific findings. However, perhaps even more important is the fact that, on the same day, Gauss decided to devote himself to mathematics instead of philology, the study of language, a field in which he had also shown flashes of genius.

Nowadays, this result is known as the 'theorem of prime numbers' and is one of the most important in the history of mathematics. The wild collection of prime numbers was beginning to be tamed. A function had been introduced to study them and as time passed it would yield ever more accurate results.

Gauss did not live to see these successes. This was not due to secretiveness, as is often the case, or to the attitude shown by Fermat, who would have said that he did not include the proof because it was too long; conjectures in those days were usually supported by piles of laboriously gathered evidence. However, we can presume that Gauss did have enough paper for any proof, however long it was. Gauss did not see the success of this theorem simply because he had no means of proving it. Mathematics had turned a corner with the work of Euler, where theories were expressed in a logical way that was beginning to blaze a trail through the ambiguous techniques and dubious practices of the past. Intuition, the key to any discovery, would have to be backed up with a solid theoretical basis. Proving a theorem had become an objective argument, which, thanks to the common language of numbers, would acquire the status of a truth.

Gauss's conjecture only became a theorem a century later. In 1896, Jacques Hadamard (1865–1963) and Charles Jean de la Vallée Poussin (1866–1962) proved the theorem simultaneously but independently, and so the recognition has gone to both. Of the many theorems devised for prime numbers, the one that Gauss started with his conjecture occupies a special place in the history of mathematics, not only for its beauty but also for the huge importance it has had in the subsequent course of research into prime numbers.

Gauss appeared on the front of the old 10 Deutsche Mark note next to the curve known as a Gaussian bell curve. The reverse depicts a sextant, the instrument used to establish one of the first geodesic networks in the world, near Hamburg, as shown in the bottom right corner. The idea of a 'geodesic', or shortest line connecting two points on a given surface, is a key concept in geometry and is another of the scientific contributions made by this amazing German genius.

Chapter 5

Cornerstones

There are three theoretical developments that form the foundations of the modern study of prime numbers: modular arithmetic, complex numbers and the analytical theory of functions. The last of these is the one requiring the most mathematical knowledge if it is to be successfully applied. However, there is an aspect of it – the attempt to visualise a function using a four-dimensional space – that can be easily understood, and it will help us appreciate how Riemann's zeta function finally managed to impose order on the chaotic sequence of prime numbers.

Magic sums

As is well known, numbers have a more or less precise symbolism that takes different forms according to the mystical thinking associated with them. Most of these symbols, at least in the Western world, have a common source in the Bible and also in the Pythagorean school. "Everything that can be imagined has a number, as it is impossible for anything to be conceived or known without a number," stated Philolaus of Croton, (c. 480 BC), a Greek mathematician and philosopher who was a disciple of Pythagoras.

Transmission of the 'number culture' slowed to a trickle during the dark years of the Middle Ages. The Catholic Church made a clear distinction between various philosophical concepts of the world and those unchallengeable principles that conformed to its doctrine. One tradition that succeeded to a certain extent in penetrating this intolerance was tarot. Although the Church also ended up condemning it, the numerology of tarot was preserved in many texts that are so ambiguous that it is unclear whether they record fortune telling or arithmetic.

Based on a decimal system of counting, tarot accorded specific significance to the first nine numbers. The number 1 is unity and represents uniqueness; 2 is a symbol of difference and reproduction; 3 is the direction taken by 2 with the addition 2 + 1. Taking a similar example, 7 represents the result of the potential contained in 6: 7 = 6 + 1. And so on.

Thus, starting with 1, the basic principles of the first nine numbers are established, and it should be possible to reduce every other number to one of these. This, then, is where the notion of a 'magic sum' comes in. The basic idea is to add up all the figures in the number in question and so reduce them to a single figure. For example, take the number 47. This is reduced as follows: $4 + 7 = 11 = 1 + 1 = 2$. Hence, the number 47 inherits the symbolism of the number 2 but is situated on a higher plane. Another example of this would be:

$$157 = 1 + 5 + 7 = 13 = 1 + 3 = 4.$$

The basic operations of addition and multiplication could also be carried out by reduction. Hence, to add the numbers 248 and 386 we first carry out the reductions

$$248 = 2 + 4 + 8 = 14 = 1 + 4 = 5$$
$$396 = 3 + 9 + 6 = 18 = 1 + 8 = 9,$$

so that the sum of these two numbers would be:

$$9 + 5 = 14 = 1 + 4 = 5.$$

If, instead, we carry out the addition first and then reduce the result, we get the following:

$$248 + 396 = 644 = 6 + 4 + 4 = 14 = 1 + 4 = 5.$$

NUMBERS AND LETTERS

In the Greek and Hebrew cultures, the letters of the alphabet were also associated with numbers, so that words could acquire different mystical meanings. The basic process involved adding the numbers associated with each letter. To compare two words, the corresponding numbers were compared, and the one giving the larger number was considered more important. Legend relates that the superiority of Achilles over Hector was due to this calculation: the word Achilles adds up to 1,276, whereas Hector's result is only 1,125.

Hence, the same result of the addition is obtained when the operation is carried out in a different order.

For a multiplication we proceed in a similar way:

$$45 \cdot 27 = 1{,}215 = 1 + 2 + 1 + 5 = 9$$
$$45 = 4 + 5 = 9$$
$$27 = 2 + 7 = 9$$
$$9 \cdot 9 = 81 = 8 + 1 = 9.$$

We can arrange the first 100 natural numbers in columns, so that each column contains all equivalent numbers, according to the above reduction system:

1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18
19	20	21	22	23	24	25	26	27
28	29	30	31	32	33	34	35	36
37	38	39	40	41	42	43	44	45
46	47	48	49	50	51	52	53	54
55	56	57	58	59	60	61	62	63
64	65	66	67	68	69	70	71	72
73	74	75	76	77	78	79	80	81
82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99
100	...	...	...	...	...	...	...	...

We can now say that 78 belongs to the group of 6 or that 93 to that of 3. In the language of modern mathematics, these groups are called 'equivalence classes'. Thus we can speak of the 'class of 3', the 'class of 5', etc.

This type of classification, already known to mathematicians of his day, led Gauss to devise a new calculating tool that turned out to be very useful for determining some of the characteristics of prime numbers.

MAGIC SQUARES

Magic sums were usually additions performed in magic squares – an arrangement of numbers in the form of a square, so that adding rows, columns or diagonals always adds up to the same number, called the magic constant. Magic Squares have a long history, dating back to 650 BC in China, and most cultures have developed them. Many famous mathematicians, such as Stifel, Fermat, Pascal, Leibnitz and even Euler, were interested in this kind of arrangement. Nowadays, algorithms have been devised to construct most magic squares.

A magic square contained in the painting Melancholy I, *a work by the Renaissance artist Albert Dürer.*

Gauss's clock

The face of a clock has 12 numbers arranged around the edge of a circle. After 12 should come 13, but what we actually do is go back to the beginning and start counting again. This system is practically the same as the one described for magic sums, except that, instead of counting again from 9 onwards, we go up to 12. We could draw up a table similar to the previous one, with 12 columns in place of nine. We set out just the first two rows of this table:

1	2	3	4	5	6	7	8	9	10	11	12
13	14	15	16	17	18	19	20	21	22	23	24
...	...	...	...	...	...	...	...	...	...	...	...

This is precisely what we do every time we look at a 24-hour clock. To distinguish the hours before noon from those that come after it, we count up to 12 and then start again from 1. For example, when we refer to 17:00 hours, we know this is equivalent to '5 in the afternoon' – 17 is in the same 'class' as 5, according to our table. This gave Gauss the idea of working with different clocks or, more precisely, different clock faces. For example, a clock with only five hours marked on it would yield a table like this:

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
...	...	...	...	...

Thus, according to our previous criterion, we can say that the number 17 is in the group of 2 or, more properly, it belongs to the 'class of 2'.

It is easy to tell which class a number belongs to. Take the example of 18. We would have to go three times around the clock to reach 15 and then start again to get to 3, meaning that 18 belongs to the class of 3. That is the same as dividing 18 by 5 with a remainder of 3. This operation is very useful for large numbers. If we want to know to what class the number 40,248 belongs to, for example, we obtain a quotient of 8,049 with a remainder of 3; therefore, it belongs to the class of 3. As multiples of 5 all give a remainder of 0, we use 0 to denote the class of 5, and the above table is rearranged to become:

0	1	2	3	4
10	6	7	8	9
15	11	12	13	14
20	16	17	18	19
...	...	...	...	...

We could say that 17 is the same as 2, but such an equality expressed as $17 = 2$ would be too confusing, and so we usually write it as $17 \equiv 2$.

We might agree that an expression of this kind is correct but something is missing. We need to know what type of 'clock' we are dealing with. In the case in point it is a clock with just five digits on its face, and we indicate this by writing mod. 5, and the above expression then becomes:

$$17 \equiv 2 \pmod 5.$$

This expression is the same as saying that 17 and 2 are equivalent in modulo 5. As was usual in his day, Gauss used Latin for his scientific writing and this is why he chose the word *modulo* (the ablative of *modulus*, which means 'absolute value'). In this way, what we now call modular arithmetic was born and is now one of the most powerful tools in number theory.

Identities

Modular arithmetic uses identities instead of equalities, so the correct way to refer to the above expression is '17 is congruent with 2 modulo 5'. To find out if any two numbers are congruent in modulo 5, we subtract one from the other and see if the result is a multiple of 5. In the above case we would have $17 - 2 = 15$, which is a multiple of 5.

$82 \equiv 58 \pmod 4$ because $82 - 58 = 24$, which is a multiple of 4.

Once we have decided on a modulus (a face on one of Gauss's clocks) we can speak of its groups or classes. Suppose that we select a face with four numbers, that is, we work with modulus 4. We will have only four groups or classes of number, and only the simplest of members: 0, 1, 2 and 3. This means that we can use the number 2 in place of 382, as 382 divided by 4 gives a remainder of 2. This means that we can draw up the addition table as follows:

0	1	2	3
1	2	3	0
2	3	0	1
3	0	1	2

To reiterate, $2 + 3 = 5$, but on a clock face with four numbers, 5 would be equivalent to 1 or $5 \equiv 1 \pmod 4$. Following that same pattern, the multiplication table would be:

1	2	3
2	0	2
3	2	1

This table contains the curious situation in which two numbers other than 0 multiply together to give 0 ($2 \cdot 2 = 0$). The same would happen with the numbers 2 and 3 if we constructed a modulus 6 multiplication table, as multiplied together they would give 6, which is the same as 0, as $6 \equiv 0 \pmod 6$. This does not happen if the number we take as a modulus is prime, because a prime cannot be broken down into a product of factors.

This is where prime numbers come into their own. Congruencies are studied at secondary school and are, to a certain extent, a pleasant enough topic, but it is when we turn to the intricacies of modular arithmetic that things get really interesting, and prime numbers become indispensable.

The 'clock calculator' created by Gauss turned out to be an extremely powerful tool. He could find out, for example, that the result of dividing 8^{514} by 7 gave a remainder of 1 without having to do complicated calculations, as $8 \equiv 1 \pmod 7$ or, equivalently, 8 divided by 7 gives a remainder of 1, which in the multiplication table means that multiplying 8 by 8 is the same as multiplying 1 by 1:

$8 \cdot 8 = 64$, which divided by 7 gives a remainder of 1.

Consequently, multiplying 8 by itself 514 times is like multiplying it by 1 the same number of times; expressed in another way,:

$$8^{514} \equiv 1 \pmod 7.$$

Gauss noticed on his clock calculator that when the face contained a prime number of hours, p, these were repeated every pth time, that is, they formed iterative cycles of p numbers. Gauss then reformulated Fermat's little theorem in terms of the clock calculator, as follows:

"If p is a prime number, then for every natural number a, $a^p = a \pmod p$".

Or, equivalently, $a^p - a$ is a multiple of p. For example, $3^5 - 3 = 240$, which is a multiple of 5. In the terms of Gauss's clock, the theorem can be interpreted as follows. Suppose we wish to know whether p is a prime number. We construct a clock with a face showing p hours, take different numbers and check whether the hands point to the same digits when we raise p to each of these numbers. If not, we can be sure that it is not a prime number. Suppose that the number we wish to investigate is 6. We

construct a clock with six hours. We now take one of these hours, for example 4. We write $4^6 = 4{,}096$ which, divided by 6, gives us a remainder of 4.

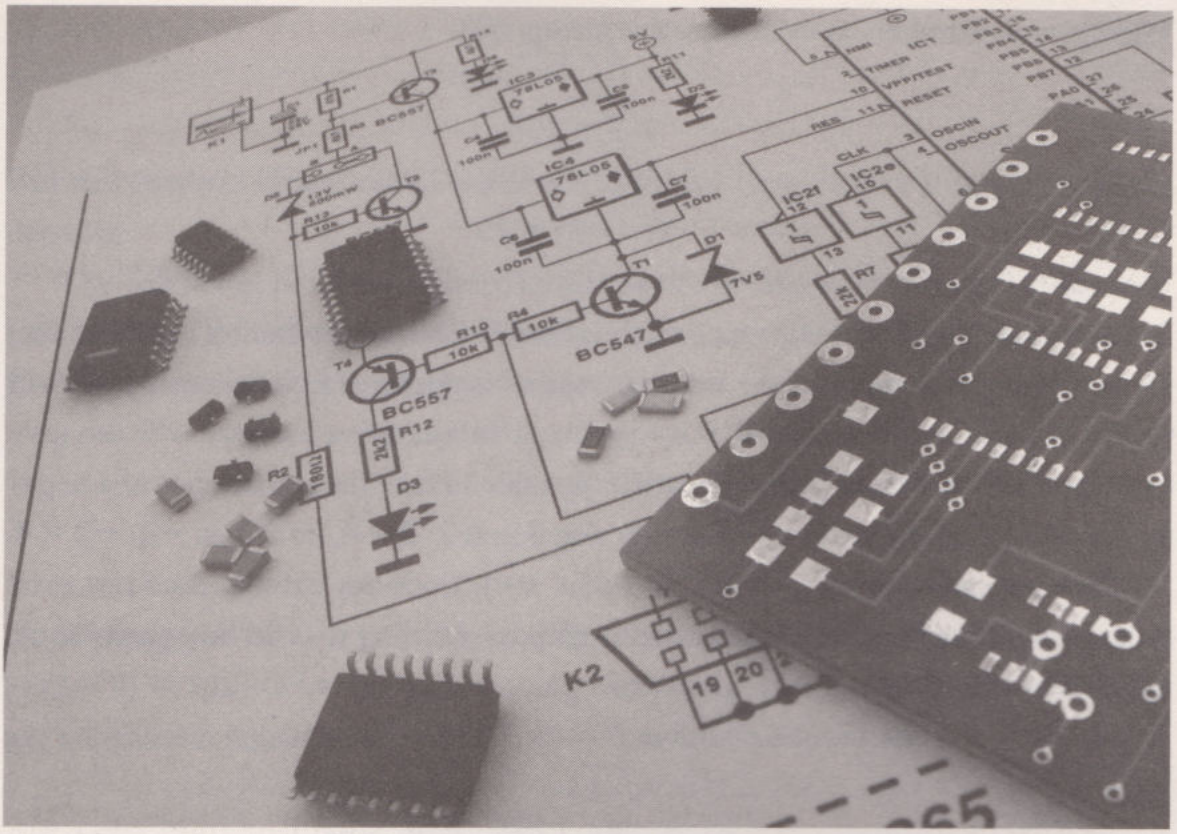
Stated another way, the hands go round and round the clock face and stop at the number 4. We know from Fermat's little theorem that 6 is not a prime number. We can then try this with a prime number, for example 7, and see that, when we raise this number to any number, the hands always return to the same hour. However, we need to bear in mind that the theorem is based on a necessary but insufficient condition. This means that, if, when we test this with a, the hands return to the number a, we know there is a good chance that the number p is prime. However, the test is not conclusive. The more tests we make, the more certain we will be that the number in question is prime but we cannot say so for sure. As we shall see in Chapter 7, this is one of the systems widely used by modern computers to find out whether a large number is likely to be prime.

Imaginary numbers

On hearing the expression 'imaginary numbers', the layman might think this is just another extravagance perpetrated by mathematicians – a not altogether unreasonable assumption. It was an opinion long shared by many professionals in the mathematics community, who occasionally encountered numbers that were so exotic that they were treated almost as invisible ghosts. But these ghosts kept turning up in the solutions of equations, and it became very difficult to ignore them. They began to be introduced into calculations until, eventually, they were accepted as solutions to equations and acquired their own identity, becoming a fundamental concept in mathematics and an essential topic in any textbook. It would be wrong to suppose that their presence is limited to the world of pure mathematical theory. In fact, imaginary numbers are a basic tool of modern physics and they have innumerable practical applications.

If logarithms played a decisive role in the contribution Gauss made to the history of prime numbers, imaginary numbers would be indispensable to Riemann's later theories, and so a short journey into the land of the 'imaginary' is essential for a better understanding of the revolution that his theory fomented.

Gottfried Leibniz once said: "The divine spirit found a sublime expression in this marvel of analysis, this portent of an ideal world, this hybrid of existence and non-existence which we call the imaginary root of minus one." We now examine what is meant by the 'imaginary root' of -1.



Imaginary numbers have a practical application in electronics. Real numbers are used to measure resistance – the opposition shown by an object when an electric current passes through it. However, imaginary numbers are used to measure inductance, the ratio of magnetic flux to current strength in a coil, and the capacitance, the difference in voltage between the plates of a condenser and the electric charge stored in it.

The square root of a number a, written as $\sqrt{a}$, is by definition another number b such that, when squared results in a. In other words, $\sqrt{a} = b$ means that $b^2 = a$. For example:

$$\sqrt{4} = 2 \text{ because } 2^2 = 4$$
$$\sqrt{9} = 3 \text{ because } 3^2 = 9$$

It should be noted that there is a 'rule of signs' in multiplication and division. A plus times a plus equals a plus; a plus times a minus equals a minus; and a minus times a minus equals a plus.

When written in symbols this becomes:

$$+ \times + = +$$
$$+ \times - = - \times + = -$$
$$- \times - = +.$$

Taking some numbers as examples, this means that:

$$5 \cdot 2 = 10$$
$$(-5) \cdot 2 = -10$$
$$(-5) \cdot (-5) = 25.$$

Thus, the square of a number, which is the number multiplied by itself, can never give a negative result. If the starting number is positive, 'a plus times a plus' will yield a positive, and if the number is negative, a 'minus times a minus' will also give a positive. This is why, in principle, it is impossible to take the square root of a negative number. For example, $\sqrt{-4}$ cannot equal 2, as $2 \cdot 2 = 4$, or -2, as $-2 \cdot -2 = 4$.

Thus we can state that $\sqrt{1} = 1$, but that $\sqrt{-1}$ does not exist. It does not exist as a real number but there is nothing to stop us defining it as an 'imaginary' one, which we will call i:

$$\sqrt{-1} = i.$$

Let's see what happens to this new number we have obtained when we raise it to different powers:

$$\sqrt{-1} = i;$$
$$i^2 = \left(\sqrt{-1}\right)^2 = -1;$$

$$i^3 = i^2 \cdot i = (-1) \cdot i = -i;$$
$$i^4 = i \cdot i^3 = i \cdot (-i) = -i^2 = -(-1) = 1.$$

And, following this pattern, we can continue thus:

$$i^5 = i;$$
$$i^6 = -1;$$
$$i^7 = -i;$$
$$i^8 = 1;$$
$$\ldots$$

The need to find the value of the square root of negative numbers arises when we solve certain quadratic equations. It was known that an equation of the type $ax^2 + bx + c = 0$ has two solutions given by the formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

But this method always failed when the quantity appearing within the root was negative.

The following problem appears in the work *Ars Magna* by Girolamo Cardano (1501–1576), published in 1545: "Divide 10 into two parts whose product is 40." If we call these two parts x and y, we can state:

$$x + y = 10;$$
$$x \cdot y = 40.$$

Writing $y = 10 - x$ and substituting into the second equation, we arrive at: $x(10 - x) = 10x - x^2 = 40$, and, moving everything over to the left-hand side, we get a quadratic equation, $x^2 - 10x + 40 = 0$, whose solutions are:

$$x = \frac{10 \pm \sqrt{100 - 160}}{20} = \frac{10 \pm \sqrt{-60}}{20} = 5 \pm \sqrt{-15}.$$

Cardano studied the two numbers he obtained as solutions:

$$5 + \sqrt{-15} \text{ and } 5 - \sqrt{-15}.$$

Aware that these are complex numbers, he noted that the sum is 10 and the product is 40 and that, therefore, despite "the mental torture this implies", they were solutions to the stated equation.

These 'complex' roots frequently appeared as solutions in many problems. (When we speak of the roots of an equation we are referring to its possible solutions.) They existed and troubled mathematicians, who could not accept them as numbers. Descartes himself said: "Neither true roots nor false roots are always real, sometimes they are imaginary," thereby coining one of the terms used ever since to denote this kind of root: 'imaginary'.

An imaginary number like $\sqrt{-4}$ may also be written as $\sqrt{4} \cdot \sqrt{-1} = 2 \cdot \sqrt{-1}$ and as we have called i the square root of -1, we can also say that:

$$\sqrt{-4} = 2i.$$

Hence, any complex number can be written in the form $a + bi$, called the *binomial form* of a complex number, in which a is the real part and b the imaginary part. For example, the number $2 + \sqrt{-9}$ can also be written $2 + 3i$, with 2 as the real part and $3i$ as the imaginary part. When a complex number has no real part, like $2i$, it is called a pure imaginary number.

Adding and subtracting complex numbers is very easy. The sum of two complex numbers is another complex number whose real part is the sum of the real part of each of the numbers and whose imaginary part is the sum of the corresponding imaginary parts.

For example:

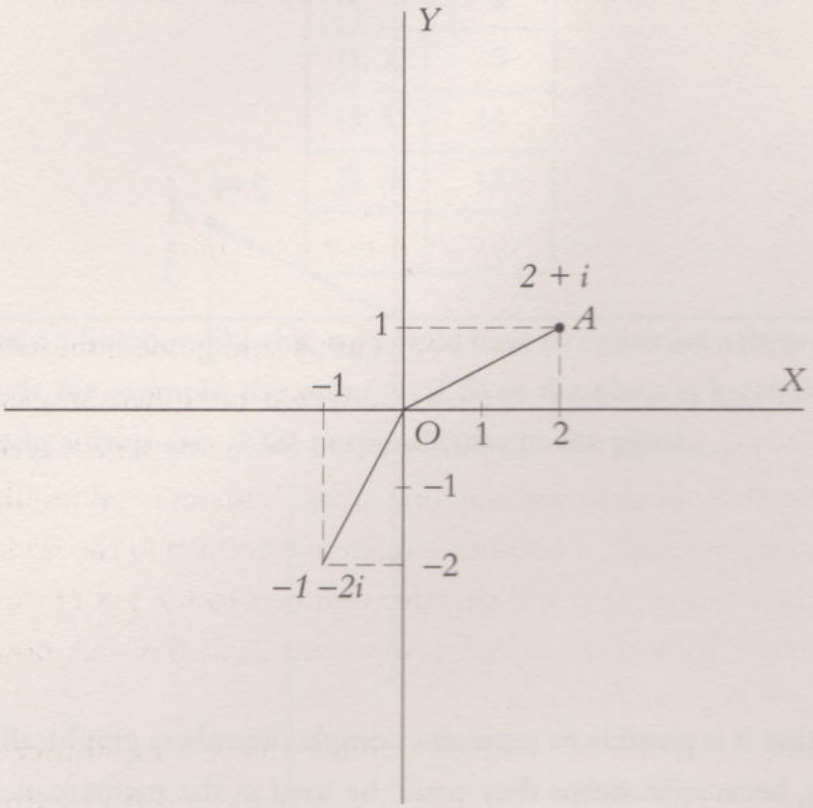
$$(3 + 2i) + (8 - 3i) = (3 + 8) + (2 - 3)i = 11 - i.$$

Subtraction follows a similar rule. For multiplication, one number is placed beneath the other and we multiply them as we would any other numbers.

In algebraic terms, complex numbers could be manipulated freely, but it was difficult to visualise them in the way one would visualise real numbers that can be arranged along a straight line, with a reference point 0 in the middle and with positive numbers to the right and negative numbers to the left of this origin. But complex numbers are written as two components, and this somehow implied an extra dimension in geometric space.

The graphic depiction of complex numbers has a long history. Several mathematicians, especially Euler, Abraham De Moivre and Alexandre-Théophile Vandermonde, had already considered the possibility of imagining a complex number $x + yi$ as a point in a plane with coordinates (x, y). However, it was Jean Robert Argand (1768–1822), an accountant and amateur mathematician, who contributed a short study on how to represent complex numbers geometrically, and a further development by Gauss who determined their geometric nature, which provided the final form that we use today. In fact, it was Gauss who introduced the symbol i to represent $\sqrt{-1}$, and he thought that $1, -1, \sqrt{-1}$ should not be considered merely positive, negative and imaginary but also forwards, backwards and sideways forms of the number 1. Indeed, imaginary numbers would have been accepted more quickly if the air of mystery surrounding them had been dispelled. For the same reason, Gauss introduced the term 'complex number' in place of 'imaginary number'.

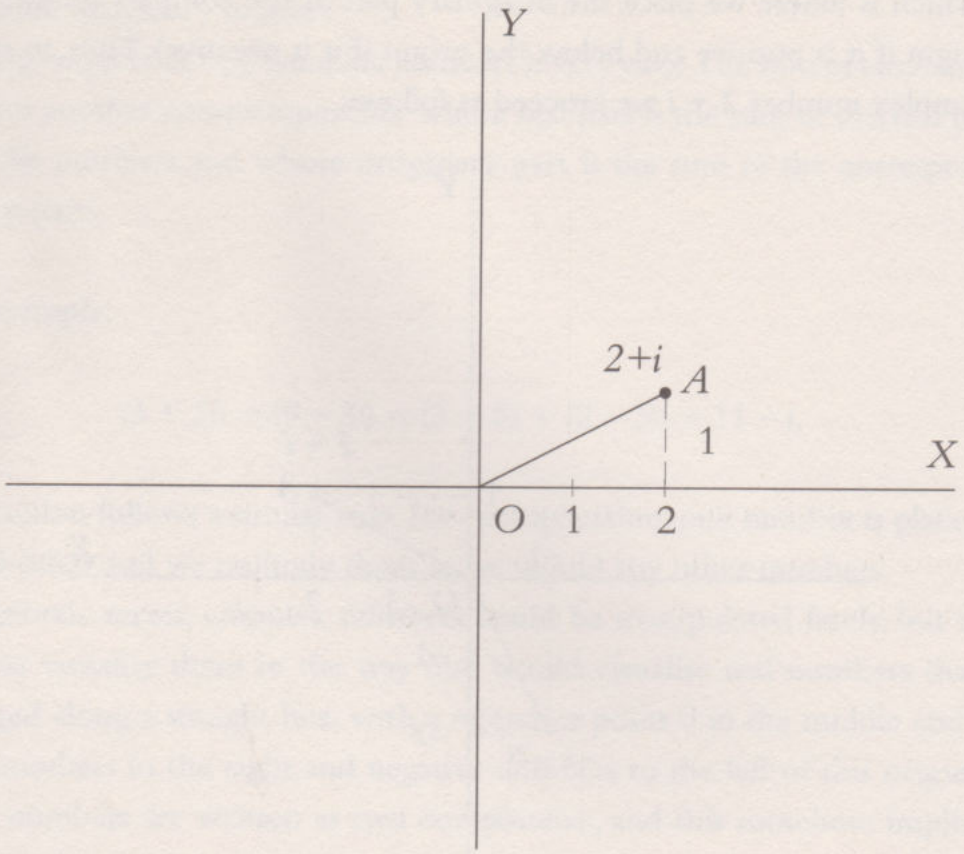
Representing complex numbers is simple. We construct a plane with perpendicular coordinate axes. We call the horizontal *OX* axis the 'real axis', which is where we place the real part of the complex number (to the right of the origin if it is positive and to the left if it is negative). We call the vertical *OY* axis the imaginary axis, which is where we place the imaginary part of the complex number (above the origin if it is positive and below the origin if it is negative). Thus, to represent the complex number $2 + i$ we proceed as follows:



FUNCTIONS INVOLVING COMPLEX NUMBERS

Ever since Cardano made his first calculations using imaginary numbers at the beginning of the 18th century, mathematicians tried to avoid any contact with numbers the existence of which they seriously doubted. Mathematicians of the stature of Euler, Wallis or D'Alembert coped with them with varying degrees of success. Complex numbers began to find uses in specific contexts, especially at intermediate stages of some proofs. Gauss was one of the first to show real familiarity with them and even devised a way of depicting them, but they were not firmly established in mathematics until the 19th century, when Riemann introduced complex functions, functions $f(x)$ in which the variable x is a complex number.

We move two units along the positive part of the OX axis and one unit up the OY axis. We can calculate the distance OA using Pythagoras's theorem, whereby $(OA)^2 = 1^2 + 2^2 = 1 + 4 = 5$, hence $OA = \sqrt{5}$, a quantity known as the *modulus* of a complex number.



The fact that it is possible to represent complex numbers graphically was a great step forwards, because it meant they could be used in the mathematical analysis of functions in which the variable was a complex number.

An extra dimension

An expert eye can derive unexpected levels of information from the graphical representation of a function. Indeed, we could consider these graphics as works of art. As Lord Kelvin once said: "A single curve, drawn in the manner of the curve of prices of cotton, describes all that the ear can possibly hear. In my opinion, this is a marvellous demonstration of the power of mathematics."

We already saw in Chapter 3 that it is possible to write functions in which every real number is assigned another real number. Using a similar mechanism, we can represent functions that assign a real number to any pair of real numbers.

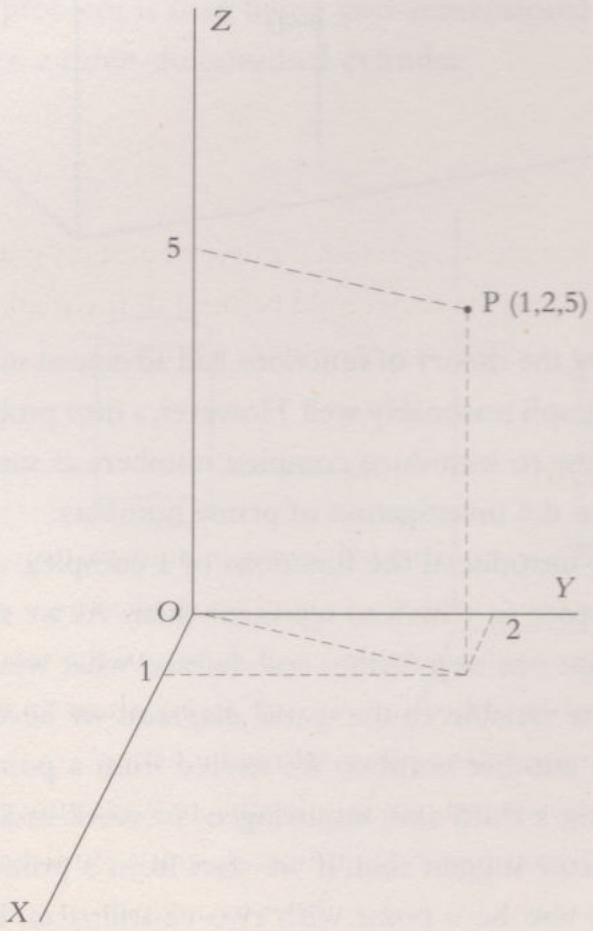
For example:

$$(x, y) \rightarrow x^2 + y^2.$$

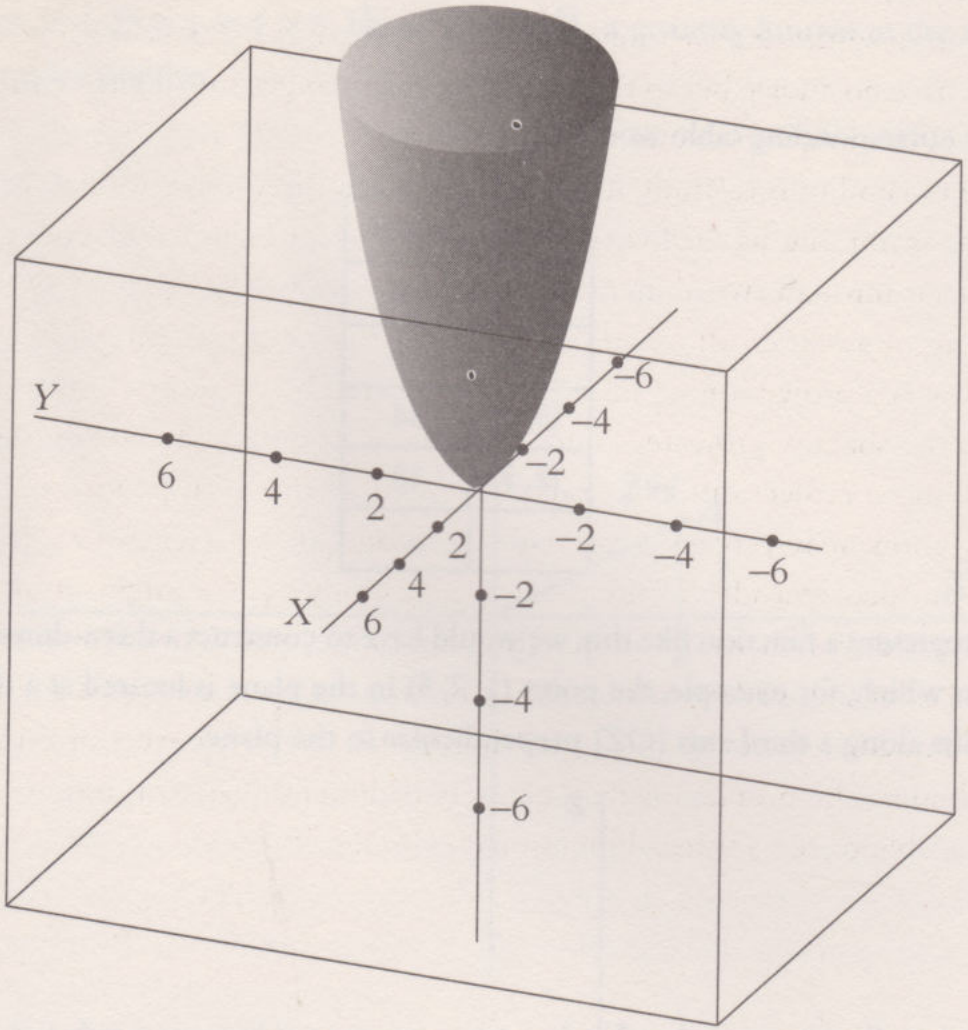
The corresponding table would be:

(x, y)	$x^2 + y^2$
(1, 1)	2
(1, 2)	5
(3, 5)	34
(2, 3)	13
...	...

To represent a function like this, we would have to construct a three-dimensional space in which, for example, the point (1, 2, 5) in the plane is located at a distance of 5 units along a third axis (OZ) perpendicular to the plane.



The function $f(x, y) = x^2 + y^2$ would be represented in the following way:

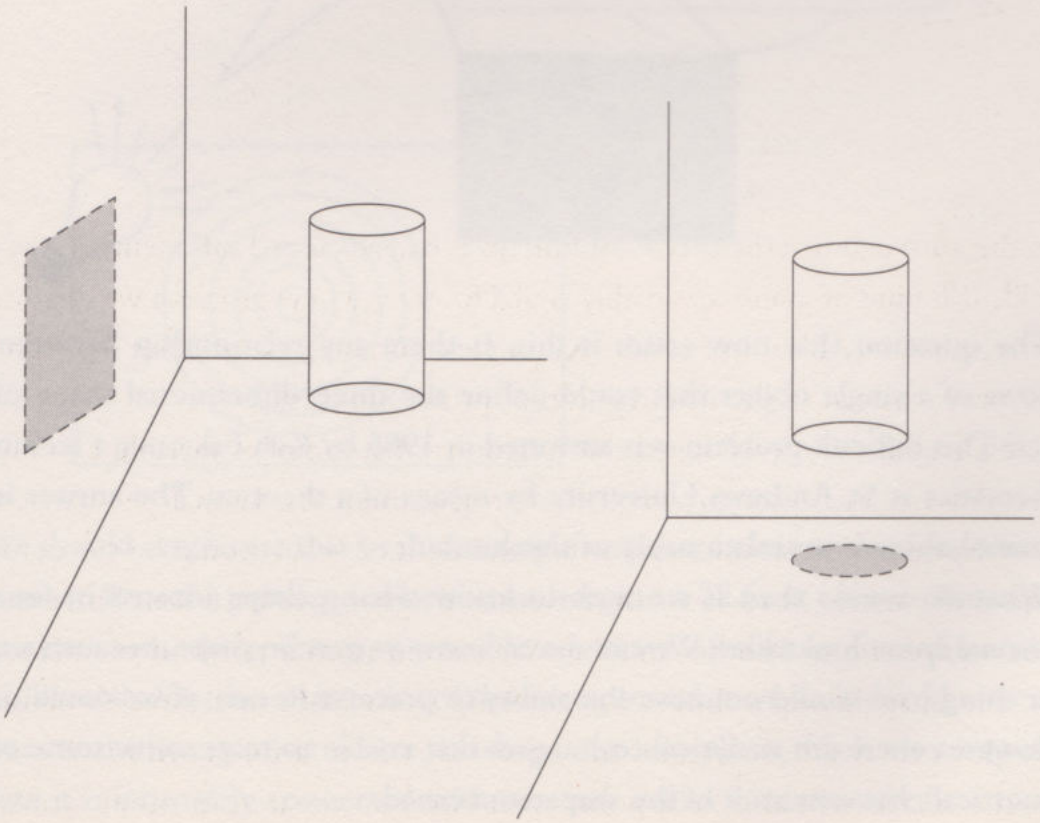


In the 19th century the theory of functions had advanced sufficiently to be able to tackle this kind of graph reasonably well. However, a new problem was beginning to emerge, namely how to introduce complex numbers as variables – a step that would prove crucial to the investigation of prime numbers.

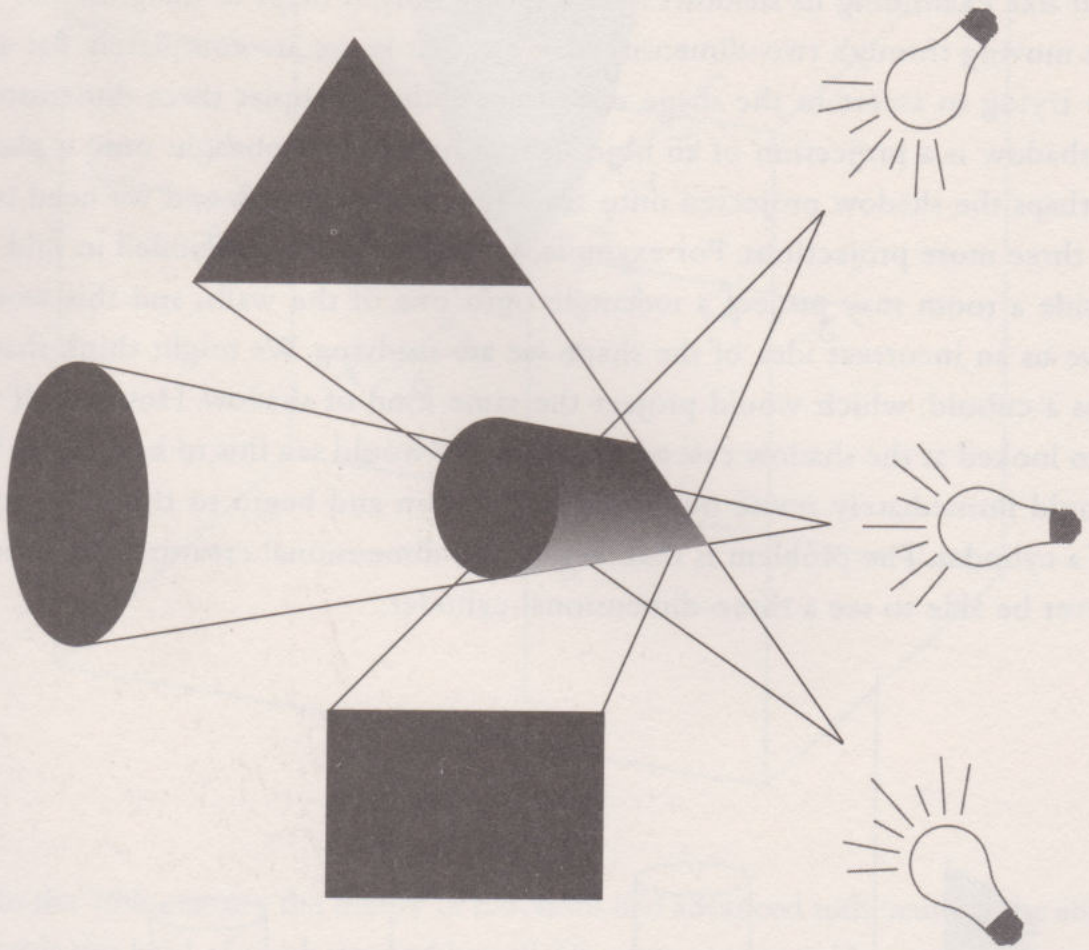
Gauss had already introduced the functions of a complex variable by drawing a three-dimensional space in which to represent them. As we shall see in the next chapter, Riemann went one step further and defined what would be the complex functions of a complex variable. In the spatial diagrams we have seen hitherto, two numbers gave rise to another number. We started from a point in the plane and traced the image along a third axis, requiring us to work in a three-dimensional space. However, we now suggest that, if we start from a point with two coordinates, the image will also be a point with two coordinates. In other words, we

need a further dimension for the graphical representation, as a function of this type can only be drawn in a four-dimensional space. Visualising an object in four dimensions is something we cannot do outside the pages of science fiction. Therefore, we have no choice but to resort to some tricks to give us an idea of the shape that the object in question may take.

One possibility is to study its projections onto three-dimensional space, rather like examining its shadows. To grasp this fully, it helps to imagine that we are moving through two-dimensional space, that is, we are completely flat and are trying to ascertain the shape of an object that occupies three dimensions. A shadow is a projection of an object illuminated by a spotlight onto a plane. Perhaps the shadow projected onto the plane is not enough and we need two or three more projections. For example, a cylinder that is suspended in mid air inside a room may project a rectangle onto one of the walls, and this would give us an incorrect idea of the shape we are studying. We might think that it was a cuboid, which would project the same kind of shadow. However, if we also looked at the shadow cast on the floor, we would see this to be a circle. We would immediately revise our initial impression and begin to think in terms of a cylinder. The problem is that, being two-dimensional creatures, we would never be able to see a three-dimensional cylinder.



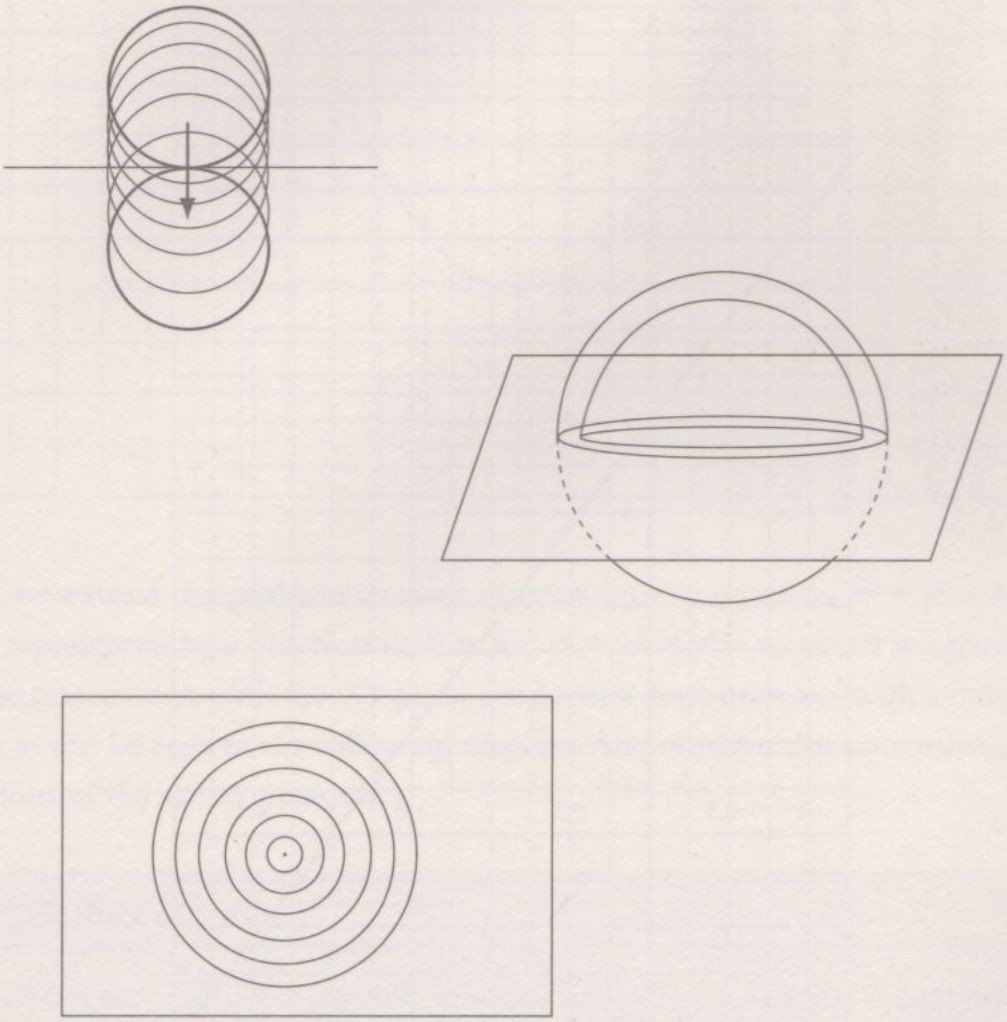
On the other hand, shadows can be very deceptive or very difficult to interpret. For example, consider an object, which when illuminated from the right casts the shadow of a circle on the wall. However, when we illuminate it from below, the shadow is triangular, and if we illuminate from above, it casts the shadow of a square. Is there any three-dimensional object with these features? If it existed, it might be a very special shape indeed!



The question that now arises is this. Is there any relationship between the shadows of a single object that could define the three-dimensional shape of the object? This difficult problem was answered in 1986 by Ken Falconer, a lecturer in mathematics at St. Andrews University, by means of a theorem. The answer is, no, in general, there is no relationship of this kind.

What do we do then if we wish to know what a shape located in four-dimensional space looks like? We can never know its precise shape, because, among other things, we would not have the ability to perceive it, even if we could depict it. However, there are analytical techniques that enable us to perceive some of the geometrical characteristics of the shape concerned.

Returning to the example in which we became two-dimensional beings, the techniques used are similar to those we would employ to determine what a sphere would look like to such beings. The trick is to see the cross sections of the sphere, where it intersects the plane we are living on – and perceiving from. When the sphere is tangential – just touching – to the plane, the first thing we would see is a point. Then a series of concentric circles would gradually appear, increasing in size and then decreasing again until they are reduced to a point when the sphere has completely passed through the plane:

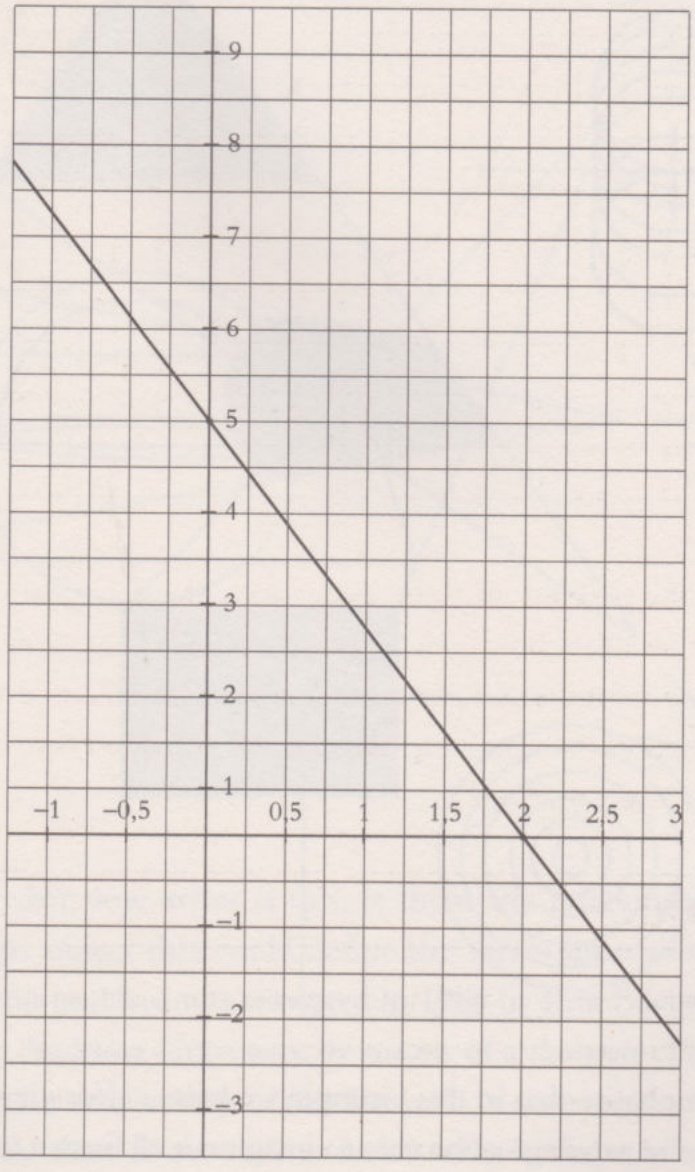


We should emphasise that in this example we have a clear view of the situation because we have the privilege of be able to imagine it all from a three-dimensional perspective, something that is impossible when we operate in four-dimensional space. Nevertheless, the main point of the example is that we do indeed know what happens at the intersection where the shape cuts our plane; this is very important because it is intimately connected with what are called the zeros of a function.

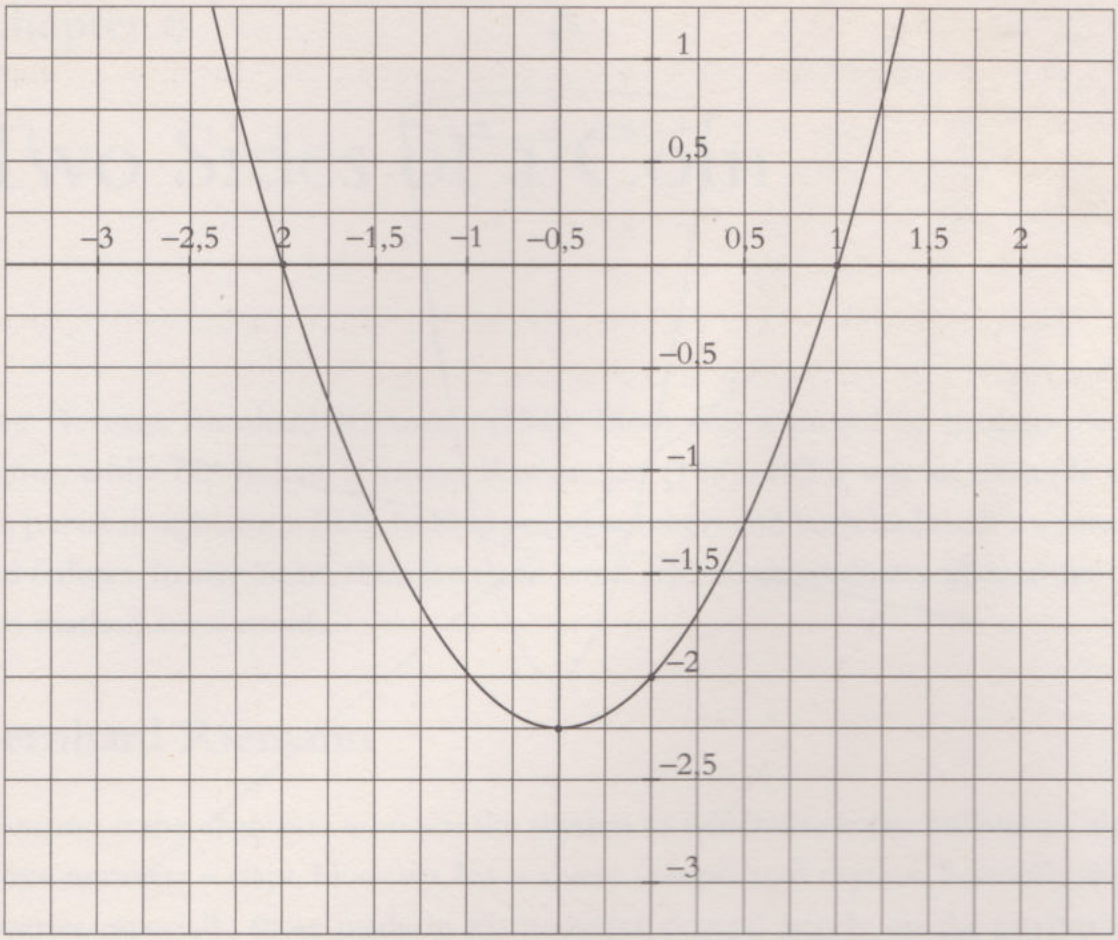
An equation such as $-\frac{5x}{2}+5=0$ can be turned into a function by simply writing:

$$y=-\frac{5x}{2}+5.$$

If we plot this we get a straight line. The intersection of this line with the horizontal axis ($x = 2$) is the solution to the first equation ($y = 0$):



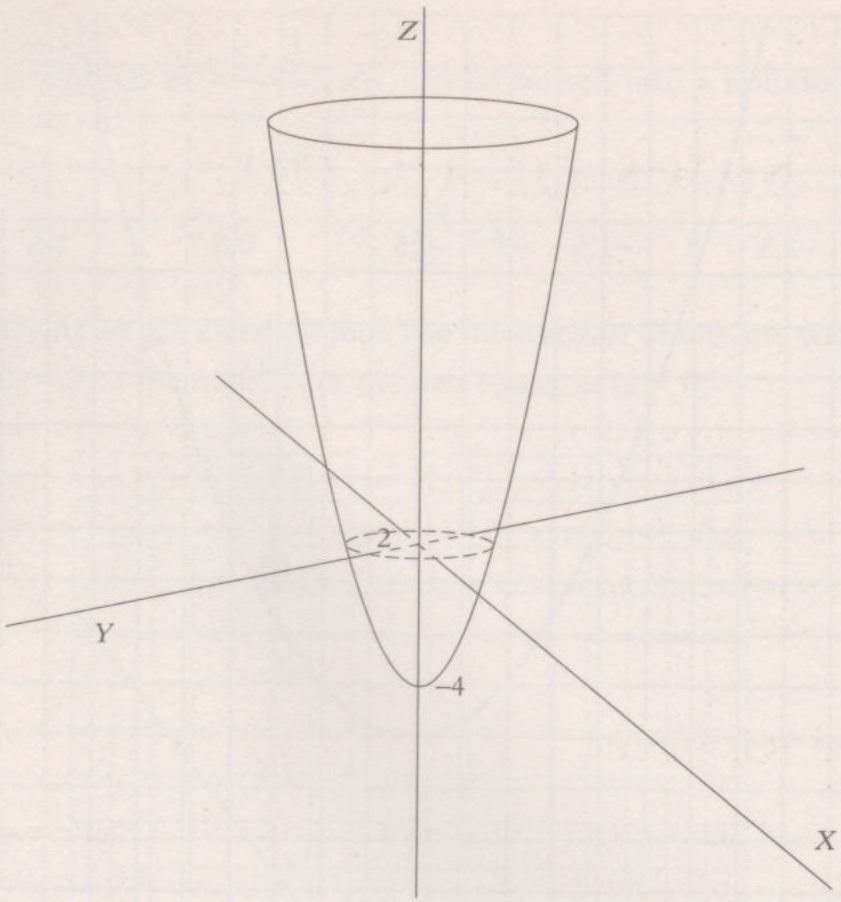
Similarly, if we have a quadratic equation $x^2 + x - 2 = 0$, and we plot the function $f(x) = x^2 + x - 2$, we will see that the function intersects the X-axis ($y = 0$) at two points, where the solutions of the equation are $x = 1$ and $x = -2$.



If we extend the problem to three dimensions, the equation $x^2 + y^2 - 4 = 0$ is then represented by a function such as $f(x, y) = x^2 + y^2 - 4$, which is a paraboloid whose intersection with the XY plane produces a circumference with a radius of 2 units, as can be seen in the following diagram. All points on this circumference are solutions of the stated equation.

CULTURAL LEGACY

If we said that 'a function is a quantity consisting in some way of a variable and of arbitrary constants,' the statement would not pass an elementary maths exam, as it would show that the person writing it did not have a clear notion of a function. However, it is taken verbatim from the writings of no less a figure than Jean Bernoulli, one of the most important mathematicians of the 18th century. In fact, it was no simple task establishing the definition of a function, as any school student will agree. This fact demonstrates the extraordinary robustness of mathematics as a cultural legacy.



Thus, when we use the trick described above to 'see' what a four-dimensional shape looks like, what we really want is to have a precise notion of how the four-dimensional figure intersects with three-dimensional space. This will not give us an exact idea of the shape – indeed we know that is impossible for us – but it will give us the solutions to the corresponding equation. And, as we shall see in the next chapter, this is what Riemann proposed when he analysed the famous zeta function that would eventually impose order on the set of prime numbers.

Chapter 6

Two Sides of a Coin

The German Bernhard Riemann (1826–1866) was a model of mathematical rigour, while the Indian Srinivasa Ramanujan (1887–1920) was an example of the purest imagination. Both tackled prime numbers and both had their successes and failures. In any event, their lives and work are an extraordinary illustration of two mathematical minds.

Bernhard Riemann

Riemann is the drummer who sets the rhythm to which the entire audience – the prime numbers – claps. However, his is a very complicated rhythm. Scientific discoveries, especially those made in mathematics, depend largely on the territories explored, on the knowledge already uncovered. The discoverer becomes something of a mountain guide. When rambling through numbers, it is important not to lose one's sense of direction, but when we start climbing them, it's a different matter. These are excursions that require more effort, and we proceed at a slower pace so that the ascent is not too tiring. However, there comes a point when hikers need a certain amount of preparation and suitable equipment to continue upward. Climbing a 2,000-m peak is not the same as climbing one 4,000 m high. With Riemann, we are definitely in the '4,000 m and above' category.

Georg Friedrich Bernhard Riemann was born in Breselenz in the Kingdom of Hanover. Perhaps because of his extreme shyness and almost pathological fear of public speaking, he did not follow the route marked out by his father, a Lutheran pastor, into the Church. Friedrich Constantin Schmalfuss, headmaster of the school where the young Riemann studied, let the boy borrow a book from his private collection, Legendre's *Theory of Numbers*, a mathematical treatise of great complexity. Riemann read it from end to end in under a week and returned it saying he had found it very interesting. He was not lying. Years later, Riemann would take what he needed from this book to develop his theory of prime numbers, thereby generating one of the most famous conjectures in the history of mathematics.

At the age of 19, Riemann attended some mathematics lectures by Moritz Stern at the University of Göttingen. This was where he first encountered the work of Gauss. A year later he enrolled for a mathematics degree at Berlin University and was taught there by Peter Gustav Lejeune Dirichlet, Carl Jacobi, Jakob Steiner and Ferdinand Eisenstein. His close relationship with the last of these teachers led to one of the most important mathematical theories of the 19th century, the 'theory of functions of a complex variable'. This became a fundamental tool that enabled Riemann to establish his hypothesis on prime numbers.

Bernhard Riemann.

DOCTORAL THESIS

"I think that I have improved my prospects with my dissertation. I also hope to learn to write more quickly and fluently, especially if I frequent [civil] society and I have the opportunity to give lectures; therefore, I am in good spirits." These words, written by Riemann in a letter to his father, referred to the examination of his doctoral thesis which, at the age of 25, he submitted to Göttingen University. Its title was *Foundations of General Theory of Functions of a Complex Variable*. It was enthusiastically received by Gauss, one of the living legends of the mathematics at the time.

The zeta function

As we saw in Chapter 3, Euler had defined a function based on a harmonic series:

$$\zeta(x) = \frac{1}{1^x} + \frac{1}{2^x} + \frac{1}{3^x} + \frac{1}{4^x} + \ldots + \frac{1}{n^x} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n^x}.$$

The Swiss mathematician had already noticed that the sum is infinite if x takes values equal to or less than 1. He succeeded in calculating two values, for $x = 2$ and $x = 4$:

$$\zeta(2) = \frac{\pi^2}{6};\ \zeta(4) = \frac{\pi^2}{90}.$$

We also saw how Euler established a relationship between this function and prime numbers (the so-called Euler product). This relationship helped him and other mathematicians to prove that primes were infinite, which Euclid had previously shown using a more elementary method.

On the other hand, Gauss had conjectured, but not proved, that for large values of x,

$$\pi(x) \approx \frac{x}{lnx}.$$

Remember that $\pi(x)$ is the number of primes smaller than x.

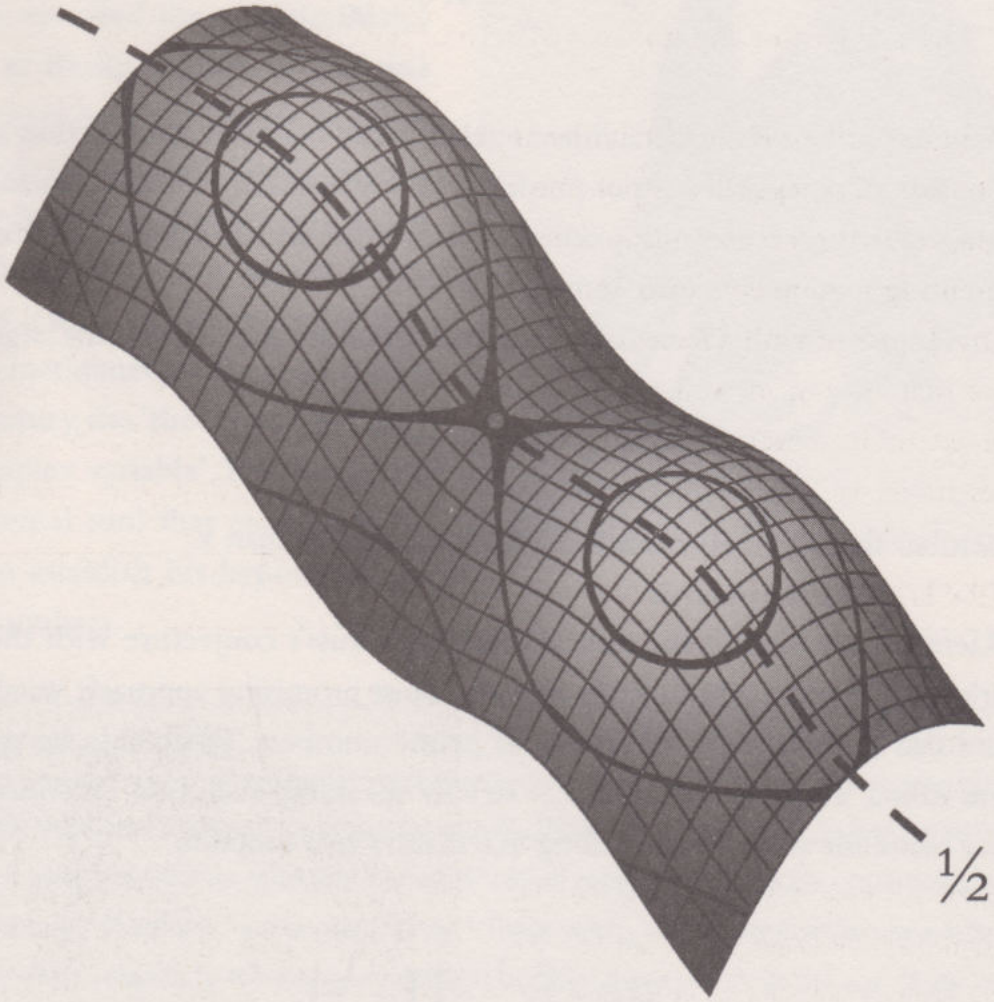
Riemann set himself the task of studying Gauss's conjecture with the aid of Euler's zeta function and thought that the most promising approach would be to extend this function into the realm of prime numbers. To do this, he devised a system called 'analytical extension' – strictly speaking, analytical extension is the proper name for what we are calling 'Riemann's zeta function':

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} = \prod_{p} \frac{1}{1 - p^{-x}}.$$

The second part of the equation, an infinite product extending to all prime numbers p, draws on the Euler product and relates the zeta function to prime numbers. Remember that this product was obtained as a direct consequence of Euclid's fundamental theory of arithmetic.

We have already seen how Gauss introduced functions of a complex variable by devising a three-dimensional space to represent them. Riemann goes one step further and defines what would later become complex functions of a complex variable. The problem was that they required a four-dimensional space and therefore could not be visualised.

Using sophisticated techniques similar to those we noted in the previous chapter, Riemann obtained a three-dimensional image of the zeros of the zeta function, a landscape consisting of hills and valleys recurring with some regularity:



In this function there are two 'classes of zero', sets of values that give zero as the result when put into the function. Some are negative even-number integers: $x = -2, x = -4, x = -6...$ known as 'trivial solutions'. There is nothing trivial about the remaining zeros, and it is extremely difficult to calculate them. There are an infinite number and all are found in a so-called 'critical strip', where the real values are located between 0 and 1 ($0 \leq \mathrm{Re}(x) \leq 1$). These form a segment of the landscape that is intimately related to prime numbers. In 1896, it was against this particular background that two mathematicians working independently, Jacques Hadamard and Charles de la Vallée Poussin, both proved the 'theory of prime numbers' proposed by Gauss.

In a somewhat informal note and without any proof, Riemann postulated that all non-trivial zeros of the zeta function were of the form $\frac{1}{2} + iy$, which was tantamount to saying that they were on the straight line $x = \frac{1}{2}$ that runs through the zeta function. This declaration forms what is known as the Riemann conjecture:

"The real part of any non-trivial zero of the zeta function is ½."

If the hypothesis is correct, it means that all prime numbers are distributed regularly, or rather, as regularly as possible. This can be understood by using an analogy. Imagine a function representing the analysis of a sound, a series of sinusoidal curves representing a violin concerto. To simplify the image, suppose that only one violin is playing. Together with a series of well-defined crests and troughs, other shapes may appear that are not so well defined and that disturb the harmony of the graph to some extent. In acoustic jargon, this is known as 'white noise' and has many possible causes: static, background sounds, etc. Now, the Riemann conjecture states that any

TRY IT FOR YOURSELF

If you wish to improve your knowledge of functions of complex variables and series – there are many excellent references available – you could try to prove the Riemann hypothesis itself. If you succeed, the Clay Mathematics Institute will award you the not inconsiderable sum of one million dollars in cash, irrespective of your age, gender or occupation. However, it will take a while for the prize to come through, as first of all the proof will have to be scrutinised and shown to be correct. In June 2004, Louis de Branges de Bourcia, a mathematician at Perdue University West Lafayette, Indiana, declared he had succeeded, but his proof was later rejected. The same thing happened in 2008 to a proof from Xian-Jin Li.

Louis de Branges de Bourcia.

irregularities appearing in the prime number distribution are due to mathematical 'white noise', meaning that the distribution of prime numbers follows a pattern that is not based on pure chance. Thus Riemann managed to impose some order on the motley crew of numbers.

In 1914, the British mathematicians Godfrey Harold Hardy (1877–1947) and John Edensor Littlewood (1885–1977) proved that there is an infinite number of zeros on the straight line. This does not prove Riemann's conjecture but it does strengthen the widespread view in the mathematical community that the hypothesis is correct. There are some who think that if there is an infinite number of zeros on the critical line, then they must all be accounted for, but this only reveals our ignorance about certain aspects of infinity, which is a concept full of paradoxes, because there may also be an infinite number of zeros that are not on this straight line. To date, around ten million non-trivial zeros positioned on the line in question have been computed.

On one occasion the eminent German mathematician David Hilbert was asked about the first thing he would ask at a mathematics symposium held 100 years after his death. He replied: "I would ask whether the Riemann conjecture has been proven." So far no one has succeeded, but Hilbert only died in 1943.

Mathematical thought

French maths genius Henri Poincaré (1854–1912) said that work in mathematics progresses in three stages. The first stage consists of a refined analysis setting out the difficulties of the problem and the different approaches needed to tackle it, the tools that are available to do so, and an acceptance that a radical rethink of our knowledge is required.

The next is a stage of apparent abandon. The mathematician stops thinking about the problem or at least stops thinking about it in a particular way, to allow the mind to explore the mysterious territory of the subconscious where creative activity obeys its own rules. This is the land of imprecision, inaccuracy and intellectual wandering. The result of this subconscious process appears as inspiration, which might come at any moment and be associated with events that have no apparent connection with the study in question. This is the moment described by the Irish mathematician Sir William Hamilton (1805–1865) when, out walking with his wife on the outskirts of Dublin, he suddenly stopped dead as if he had received an electric shock. He

continues in his own words: "That is to say, I then and there felt the galvanic circuit of thought close; and the sparks which fell from it were the fundamental equations between $i, j, k, \ldots$."

Hamilton was saying that not three but four numbers were needed to describe the spatial behaviour of a hypercomplex number. It is the magic moment when the researcher suddenly feels that a light has been switched on in a room he has never visited before.

PARADOXES OF INFINITY: THE HILBERT HOTEL

The Hilbert Hotel is an imaginary building in which there is an infinite number of rooms. Its manager boasts that he has never turned a guest away. Imagine that, one night, all the rooms are occupied and a new guest suddenly appears. The receptionist goes to the manager and informs him that the guest cannot be accommodated. The manager tells him to ask the guests who already have rooms to switch to the one next door, so that the guest in room 1 goes to room 2, the one in room 2 to room 3, etc. When this has been done, room 1 is free and is taken by the newcomer. However, at midnight, the receptionist goes to the manager again. This time he looks pretty desperate. An infinite number of mathematicians attending a symposium have just arrived. "We definitely can't accommodate them all!", he cries. After thinking for a moment, the manager suggests the following: "We will have to ask our guests for another favour. Each one must multiply the number of his room by two and move into the one with the number resulting from his calculation." That is, the person in room 4 goes to room 8, the one in room 23 to room 46, the one in room 352 to room 704, and so on. When this has been done, all the rooms with an odd number in the hotel are left empty, enabling all the symposium members to be accommodated.

Portrait of David Hilbert from 1912.

Poincaré then examines the selection process that goes on subconsciously, making us aware of certain ideas and making us reject others. Finally we reach the conclusion that because we are unable to decide whether such ideas are true or false, the only selection criterion left is mathematical beauty.

After that, in the third stage the mathematician is working fully conscious and subjects his ideas to severe scrutiny, accepting some and rejecting others. He may return once or several times to stage two until the problem has finally been solved, thus obeying the rules imposed on it by mathematical convention, and the solution takes on its finished form.

Henri Poincaré was a scientist who showed exceptional skill in all areas of mathematics.

All three stages are important in making a mathematical discovery, but for many the second stage is the most interesting because it is the one in which the mind takes flight, freed from the constraints of conscious thought. Jacques Hadamard dedicated one of his books, *Psychology of Invention in the Mathematical Field* (1945), to studying the role played by the subconscious in creative activity, concentrating mainly on the mathematical mind. He considered the creative experiences of some of the greatest thinkers of his generation, including those of Albert Einstein. Hadamard's book describes mathematical creativeness as a process that starts with deliberate choices about the most important aspects of a problem, most often after obtaining results that are inconclusive. He thought that this period should be followed by a 'rest', with the problem put aside, then followed by moments of inspiration, when the result of thought processes not perceived consciously by the mathematician would become apparent.

Lastly, Hadamard reaches the so-called 'tidying up' stage in which formalism makes its presence felt and the results are arranged sequentially. He believed

that the intervention of the subconscious was crucial throughout the process, especially in the rest period.

Hadamard's conclusions coincide with Poincaré's, except that the latter places greater emphasis on the period of rest. We should note that these periods usually include spells of sleep too. There is a lot of evidence in the history of science, and in mathematical creativity in particular, that many key ideas in research have come to light in sleep. Few researchers report that the breakthrough came in a dream in which they were working on the problem. Most say the solution came to them on waking, especially after a spell of intense work on the previous day. For example, Dirichlet used to say that he slept with Gauss's *Disquisitiones Arithmeticae* under his pillow, because he knew that, while he was dreaming, a mysterious process would occur over which he had no control and thanks to which, on the following day, he would be able to decipher obscure parts of the text that he had been unable to understand the day before.

All of this is a part of the magical world of numbers that we have entered in previous chapters. We should emphasise once again that this is not magic in the ordinary sense of the word. In their traditional sense, magical rituals or ceremonies are intended to reveal hidden truths. In the sense we are using here, a ritual, a belief or, better still, the act of dreaming itself, puts the mind in a special state, where it is free of the constraints of the physical world and can think in a different way. It is as if we had switched to another frequency band on the radio, and are able to pick up new signals, although the radio receiver used is still the same.

We store information in the brain but there can be many ways of cataloguing it. There is one mathematician in particular who exemplified the fusion of reason and imagination we have attempted to describe. Ramanujan was said to move effortlessly through the second of the three stages set out by Poincaré or Hadamard, but that he had serious difficulties with the third stage. Owing to the circumstances in which he was educated, he lacked academic training in ways to formalise his proofs according to conventions. In other words, Ramanujan could see results but he found it very hard to prove them in a way that the mathematical community considered satisfactory. Ramanujan would not become a legend, a famous maths genius in his short life, and his mathematical works were not well documented. Despite being uneducated and poor, he became one of the most important mathematicians of his day and perhaps the greatest in the history of India.

Srinivasa Ramanujan

Ramanujan was born on 22 December 1887, into a poor family in Erode, a small town some 400 km from Madras. At the age of seven, he was given a grant that let him attend classes at a school in Kumbakonam. His extraordinary ability to remember numbers and perform complex arithmetic became apparent soon after. He could repeat hundreds of decimal places of the constant π or of the square root of 2 from memory. The first mathematics textbook he encountered was G.S Carr's *Synopsis of Elementary Results in Pure Mathematics*. He was only 15, but this is when it is thought he began his first serious work in mathematics, as this was a difficult book with few proofs and should have been practically incomprehensible to him in view of his patchy mathematical background.

At 16 he obtained a grant to go to the local college in Kumbakonam. Ramanujan's passion for mathematics meant that he spent all his time on it, dropping all other subjects and so finally losing his grant. From then on he never took any subject that did not include mathematics.

In 1909 he got married and had to find a job to support his family. Through a friend he obtained a letter of recommendation to work with an amateur mathematician, Diwan Behadur R. Ramanchandra Rao, who was a tax collector in Nelore, some 130 km north of Madras. Rao described his first meeting with Ramanujan as follows:

> Some years ago, a nephew of mine who knew nothing about mathematics said to me: "Uncle, I have a visitor who talks about mathematics and I can't understand him. Can you see if there is anything of interest in what he says?" In the fullness of my mathematic wisdom, I condescended to see Ramanujan. A small, rustic, vigorous individual, unshaven, dishevelled, with an appealing face and bright eyes came in with a shabby notebook under his arm. He was extremely poor. He had migrated from Kumbakonam to Madras with the intention of finding some way of pursuing his studies. He didn't ask for

Indian postage stamp issued in 1962 to commemorate the 75th anniversary of Srinivasa Ramanujan's birth.

anything out of the ordinary. He needed to talk to someone. In other words, could I provide him with the minimum of support to allow him to think. He opened the book and began to explain some of his discoveries. I immediately realised that he was unusual but my knowledge did not permit me to judge whether or not he was talking sense. Suspending all judgement, I asked him to call again, and he did so. He was aware of my ignorance and showed me some of his simpler findings. These went further than the books available at the time, and I knew for certain that he was an extraordinary man. Then, step by step, he introduced me to elliptical integrals and hypergeometric series and, finally, his theory of divergent series, not yet disclosed to anyone – I was convinced. I asked him what he wanted. He said he wanted a small allowance to live on so that he could continue his research.

Ramanujan refused to accept charity and eventually agreed to work as an accountant for the harbour authority in Madras. Although, as a responsible worker he discharged his duties for the company, his heart and soul had but one aim, to obtain sufficient means to cover his necessities and those of his family and to devote himself to mathematics.

It is no exaggeration to say that Ramanujan had a gift for numbers. There are a some anecdotes that demonstrate his extraordinary abilities. The first of them was told by P.C. Mahalanobis (1893–1972), one of his Indian colleagues later at Cambridge University, who was trying to solve a problem of mathematical logic that had appeared in a newspaper. Having examined it for a few minutes he found the solution, a pair of numbers. He then mentioned the problem to Ramanujan, who at that moment was cooking lunch (he was a vegetarian): "Here's a problem for you..." and read it out. Instantly and without

Ramanujan's house in Kumbakonam, the town where the Indian mathematician died on 26 April 1920.

looking up from his pots and pans, Ramanujan replied: "The answer is..." And he quoted a general formula for obtaining infinite pairs of numbers, all of which were solutions to the problem. The first term was the solution that Mahalanobis had found.

The second incident occurred in the summer of 1917. Ramanujan had been interned at a sanatorium in Putney, London, with symptoms of tuberculosis. His friend and mentor, the British mathematician Hardy, went to visit him one morning. "I recall that I went to him while he was sick in Putney," Hardy related. "I had travelled in taxi number 1,729, thinking that the number was very prosaic and hoping that it did not bode ill. 'No, he replied, it is a very interesting number. It is the smallest number that can be expressed as a sum of two cubes in two different ways'. And, in fact,

$$1{,}729 = 1^3 + 12^3 = 9^3 + 10^3.$$

Of course I asked him if he knew the answer to the problem raised to the fourth power and he replied, after a moment's thought, that the example was not obvious, and the first of such numbers must be very large."

Ramanujan had been drawn into the branch of mathematics that Hardy considered the most difficult: number theory. And very soon he fell into the same trap that prime numbers had set for all mathematicians who had wandered these obscure paths through the ages. Ramanujan set himself the task of finding the 'magic formula' enabling him to predict all prime numbers. This undertaking inevitably confronted him with major problems, such as the study of divergent series.

In India his economic and social situation was preventing him from making any further progress. The mathematicians around him could not help. Some of his friends composed a letter in English, in which Ramanujan displayed his abilities and his desire to extend his knowledge, and it was sent to a number of eminent European mathematicians.

The letter read:

Dear Sir,

I take the liberty of approaching you as a clerk in the accounts department of the Port Trust Office in Madras on an annual salary of just 20 pounds. I am 23 years old. I have not had a university education but have completed my schooling. When I left school I devoted myself to the study of mathe-

GODFREY HAROLD HARDY (1877–1947)

Hardy was a colourful character with a typically British sense of humour and a very select circle of friends. One day he decided to draw up a system for evaluating people. He scored their talents on a scale of 0 to 100. He had no intention of making it public. In it he gave himself a score of 25, whereas John Littlewood got a 30 and David Hilbert scored 80 – this was Hardy's best friend and the mathematician with whom he mostly collaborated. When subjected to the system Ramanujan was awarded full marks.

According to Hardy, his greatest contribution to mathematics was his discovery of Ramanujan.

matics. I have not taken the usual route followed on a university course but am pursuing my own path. I have prepared a detailed study of divergent series in general, and the results I obtained were called 'surprising' by local mathematicians...

I would like to ask you to review the enclosed work. If you think there is something of value in it, perhaps you could publish my theorems, as I am poor. I have not included actual calculations or the expressions I used but have detailed the process I am following. In view of my lack of experience, I would be very grateful for any advice you could give me. Please forgive any inconvenience I may have caused you.

Yours faithfully,
S. Ramanujan.

Of all the many mathematicians who received Ramanujan's letter, only Hardy saw its value. Ramanujan had sent him around 120 theorems containing many formulae. Referring to these, Hardy said: "I had never seen anything like it before. Just a single page of them was enough to indicate that this could only be the work of a mathematician of the highest calibre. They had to be correct because, if they weren't, no one would have had the imagination to invent them."

In May 1913, Hardy got Ramanujan a Cambridge University grant. At first Ramanujan refused because his mother did not want him to go to England but, after a while, she relented and gave him her blessing. The reason for this, Hardy relates, was that "one morning his mother said that she had just had a dream in which her son was in a large room surrounded by Europeans, and that the goddess Namagiri had ordered her not to block her son's path in life and to help him achieve his goal."

At last, and thanks to Hardy's efforts, Ramanujan was able to study at Cambridge University, partly funded by money from Madras and partly by Trinity College. It was then that the task of the English mathematician, who would be his teacher, got really difficult. What method would he follow to teach him modern mathematics?

"The limits of his knowledge were just as vast as its depth," Hardy lamented. The difficulty was increased by the huge number of topics that Ramanujan had tackled, mixing novel results with others that had been obtained before. Ramanujan had to be largely re-educated, but Hardy tried not to break what he called the "spell of inspiration" with too much formalism.

TAXICAB NUMBERS

Since the historic meeting between Ramanujan and Hardy at the Putney sanatorium, numbers having the property of being the smallest that can be expressed as a sum of n cubes in two different ways have been called 'taxicab numbers'. They are defined thus: 'The nth taxicab number is the smallest natural number that can be expressed in n different ways as a sum of two positive cubes'. The taxicab numbers currently known are:

$$Ta(1) = 2$$

$$Ta(2) = 1,729$$

$$Ta(3) = 87,539,319$$

$$Ta(4) = 6,963,472,309,248$$

$$Ta(5) = 48,988,659,276,962,496.$$

The sixth number, Ta(6), has yet to be found.

Ramanujan (centre) and Hardy (immediate right) in a group photo taken outside Trinity College, Cambridge University.

Ramanujan lived in Cambridge for five years, during which time he published 21 articles, five of them in collaboration with Hardy who eventually declared: "I learned much more from him than he learned from me."

In the spring of 1917 the first symptoms of the tuberculosis that would eventually end Ramanujan's life appeared. That summer he entered the sanatorium. He would spend much of the rest of his life in bed. In the autumn of 1918 and coinciding with an improvement in his health, his long hoped-for appointment to a Trinity College Fellowship arrived, encouraging him to pick up his work again in what would become one of his most creative periods. In early 1919 he returned to India where he died the following year.

Most of his work is in letter form and some of it is also collected in three personal notebooks, one of which became misplaced, only resurfacing in 1976. No one has yet reviewed his work in its entirety, as, despite dying at the age of only 33, he bequeathed more than 4,000 theorems to the world of mathematics.

Ramanujan's work on prime numbers, specifically the discovery of a precise formula to obtain them, is shrouded in mystery, although to a certain extent he can be considered to have failed. Hardy noted in this regard: "Although Ramanujan had

many brilliant successes, his work on prime numbers and especially on all problems connected with this theory was certainly wrong. It could be said that this was his one great failure. However, I am not yet convinced that, in a way, his failure was not more wonderful than any of his triumphs..."

Ramanujan did not know about the work of Riemann or Gauss but tried to find a formula that would provide him with the list of prime numbers. He wanted one so that he could calculate how many primes there are below any given number. The results he sent to Hardy did not include proofs of any of his assertions. However, there was a formula that came close to spoiling Ramanujan's ambitions:

$$1+2+3+4+\ldots+\infty=\frac{1}{-12}.$$

The absurdity of this equation might lead you to suppose that its author was nothing more than a charlatan, who did not even know about convergent series. But Hardy's perspicacious mind saw in the other mathematical results supplied that his student was onto something here. The error in interpretation was corrected when they realised that there was confusion in the notation system being used. It turned out that what Ramanujan had written down was no less than one of the zeros of Riemann's zeta function, specifically the solution for $x = -1$. The method that Ramanujan said he had used had given him a formula for obtaining the number of primes between 1 and 100 million with an astonishingly small margin of error. However, Littlewood later showed that Ramanujan was wrong. Nevertheless, the search for the magic formula had led him, like many other mathematicians, into extremely fruitful areas that had a direct bearing on convergent series.

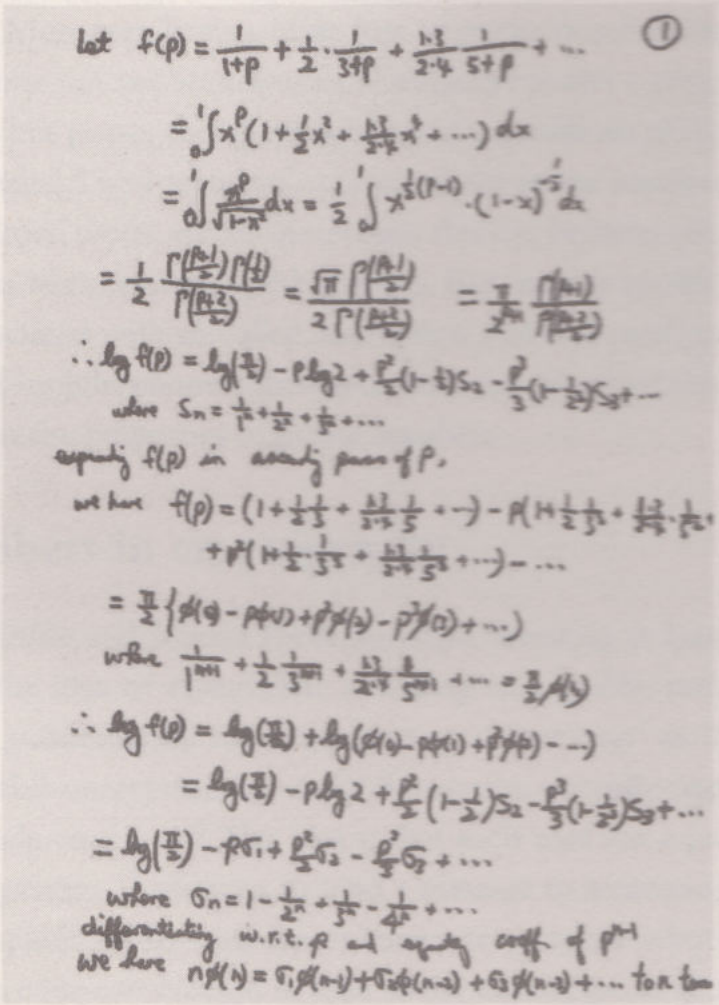
The American mathematician, Bruce Berndt, a lecturer in the maths department at Illinois University who has devoted a lot of time to studying Ramanujan's works,

AN ORDERLY LIFE

Ramanujan's lifestyle was that of a strict Brahmin, a highly spiritual caste of the Hindu system characterised by self-control and frugality and the exclusion of all animal and many vegetable products, such as garlic and onion, from the diet. It is curious to note that, throughout his life, Ramanujan wrote down many of his mathematical findings, most of which he could not yet rigorously prove, immediately after getting up in the morning.

discovered that the Indian had drawn up a table, different from the one he first sent to Hardy, in which the appearance of prime numbers among the first 100 million natural numbers is examined in greater detail. Berndt says that its accuracy is even greater than that achieved by Riemann's formula, which leads us to speculate that perhaps Ramanujan did indeed have a formula, which for some reason he kept secret. It is highly likely that Ramanujan's personal notebooks contain many more truths yet to be discovered.

It is true that Ramanujan's extraordinary mathematical mind produced some results that appear to be incorrect, but most of his results were right and possessed great mathematical beauty. In any event, his work currently occupies thousands of mathematicians in university departments across the globe, and his results are applied to areas as far removed from pure mathematics as polymer chemistry, computational design and cancer research.



A page from one of Ramanujan's notebooks.

Chapter 7

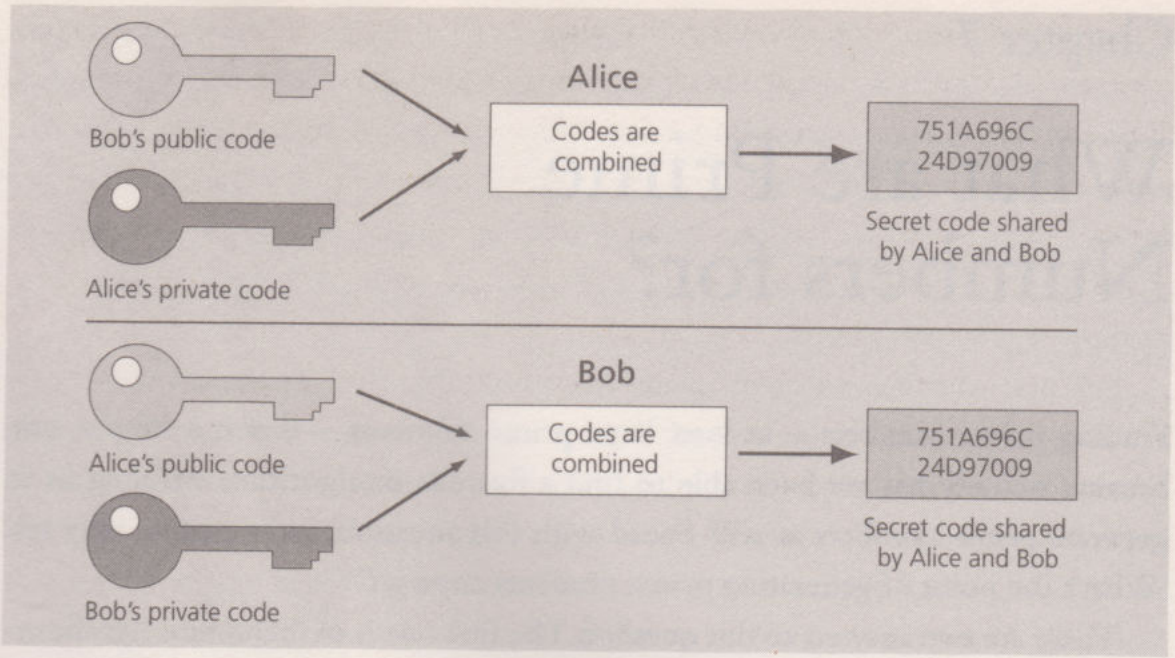
What are Prime Numbers for?

Finding prime numbers – at least large prime numbers – is not a simple task, because nobody has yet been able to find a formula or algorithm enabling us to generate prime numbers at will. Faced with this situation, some people may ask: 'What's the point of generating prime numbers anyway?'

There are two answers to this question. The first one is of theoretical significance. The drive to generate primes produces interesting calculation tools, especially for computing. Moreover, having huge lists of prime numbers helps us to check theorems that have not yet been proven. If someone makes a conjecture involving prime numbers but proves that there is one among millions of digits that violates it, the issue is settled. This has unleashed a search for prime numbers of many kinds – Mersenne primes, twins, etc. In some cases this has become distinctly competitive and involves world records and big prizes. But there is another more practical reason: one associated with so-called encryption codes. E-mail, bank transactions, credit cards and mobile phone communications are all protected by secret codes based directly on the properties of prime numbers.

Prime numbers in cryptography

In 1975, Whit Diffie and Martin Hellman, while working at Stanford University, came up with the idea of asymmetric encoding or a 'public code', a system based on specific mathematical functions known as 'one-way' or 'trap-door' functions, which enable encryption but make decryption virtually impossible without knowing the code concerned. The idea is that each user has a pair of codes, one public and one private. If we want to send a message to someone, we encrypt the message using a public code – one that is known to everyone – but only the person with the private code can decrypt it. One of the advantages of this method is that the private code is never transmitted and so does not have to be constantly replaced to maintain security. This is not a simple matter, but we can try to understand



A diagram of the theoretical principle underlying Diffie–Hellman codes. Imagine two people, Alice and Bob, need to communicate in secret. Both publicly agree on two parameters (one prime number p and another number g, with certain properties). Both Alice and Bob operate on these parameters with an integer, which they keep secret, and they publicly send each other the result of this operation. Alice and Bob manipulate this second expression and arrive at the same value, which can now serve as a shared secret code. A potential spy who has intercepted public communications between Alice and Bob could not generate the secret code from this information.

it with an analogy. Imagine a large paint shop with hundreds of thousands of cans of different coloured paints. We take two cans at random and mix the paint from each can together in different quantities. So far, so good. But if we now show the result to someone and ask him to 'decipher' the quantities of paint used in the final mixture, he would find it very hard to answer.

This is the mechanism used by mathematical trap-door or one-way functions, whereby it is very easy to 'go' in one direction but practically impossible to 'return'. Let us now suppose that, instead of paint cans, the store contains prime numbers. We take two at random, say 7 and 13, and multiply them together, just as we mixed the cans of paint, with the result being $7 \cdot 13 = 91$.

A question then arises: Is it possible to know which prime numbers have been multiplied together to give 91? This would involve having a list of prime numbers and making a few trials. It seems to be a simple matter, as indeed it would be to determine the colours in a paint mix if the store contained about only a dozen basic colours. But the reality of prime numbers is not like that.

For example, ascertaining that the number 1,409,305,684,859 is the result of

multiplying the prime numbers 705,967 and 1,996,277 would try anyone's patience, especially if we bear in mind that these two primes have been taken from a list of all prime numbers between 1 and 2,000,000, a 'mere' 148,933 of them. However, we live in a computerised age and surely this is a problem that a good program deployed in a powerful computer could solve quite quickly. Only up to a point, as it all depends on how large the paint store is, and we should not forget that the number of primes is not just very large but infinite.

The pair of prime numbers in the above example contains only a few digits. If we use primes with hundreds of digits each, the time taken by the computer program, which is looking for numbers at random – using 'brute force' as they say in cryptographic jargon – would be longer than the expected remaining life span of the Earth.

Prime numbers are completely embedded in our daily lives, in our credit cards and personal computers, and there is a demand for new prime numbers, the larger the better to construct secret codes. There is a market for prime numbers, but quality control is just as important as production. For a large number to be granted the status of a prime it needs to be tested by an officially accredited body.

The RSA cipher was published in 1978 but did not come into general use as a cryptographic code until the late 1990s, following the rise of the Internet. Finding large prime numbers was difficult, as the job required very specific software, and this tended to be bought from specialist firms or certain university departments that obtained it as a product of their research. However, the exponential growth of computing power and the constant appearance of more sophisticated implemen-

RSA 129

The RSA 129 'debacle' of April 1994 is famous in computing circles. It involved a number with 129 digits that the authors of the encryption system had made part of a public challenge. Around 600 mathematicians with the assistance of 1,600 volunteers recruited over the Internet managed to factorise the number. However, it is calculated that, if all the computers in the world were set to work in parallel, it would take the age of the universe (13.7 billion years) to break a code 1,024 digits long. Just imagine, in public-code cryptography numbers with 128, 1,024 and even 2,048 bits are used! The more digits the system has, the more resistant it will be to attack, but this also has the disadvantage of slowing down the decryption process.

tation algorithms have transformed the market for prime numbers, making them much more accessible.

The age of computing

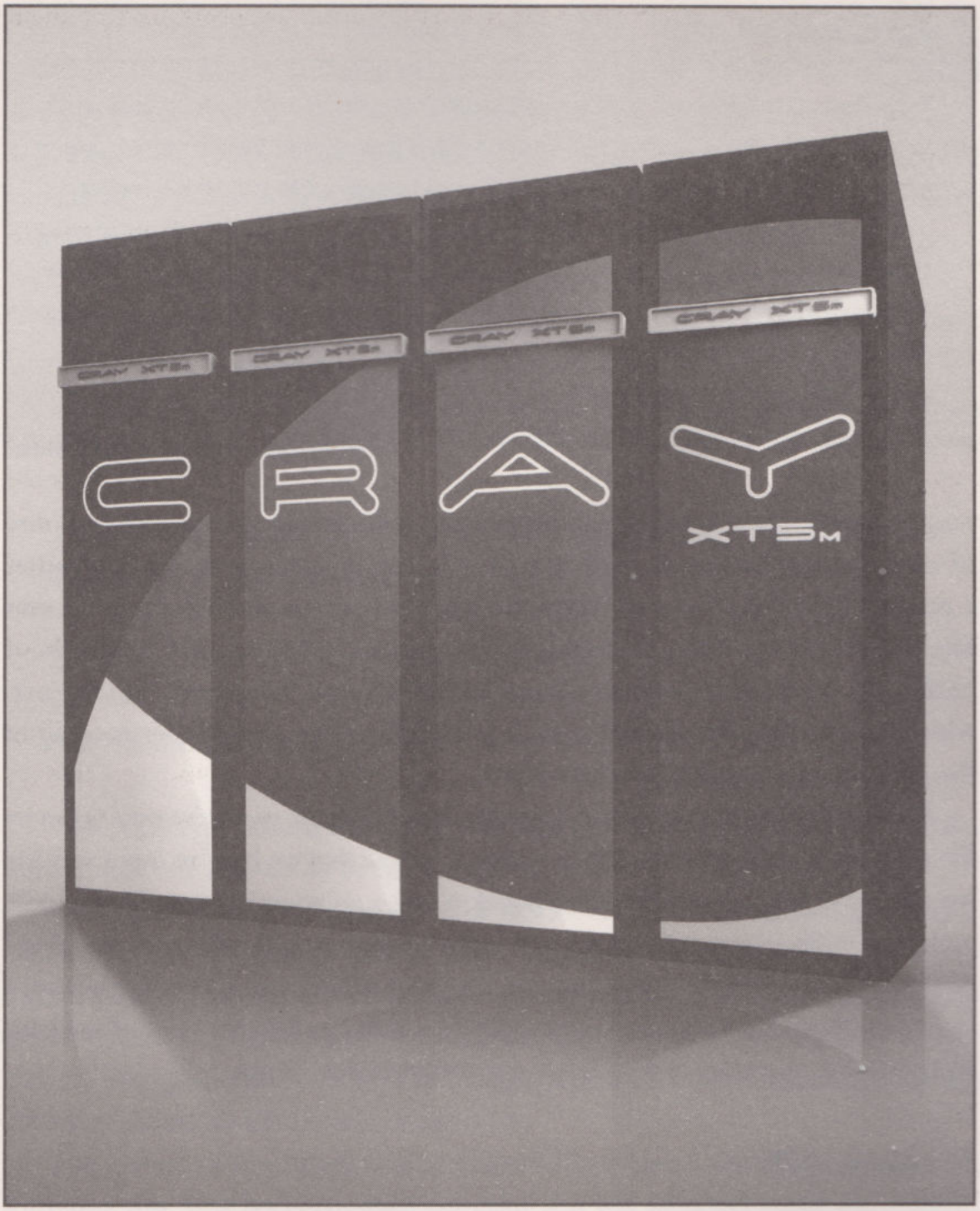
The appearance of logarithms proved a great time saver when performing calculations. Later, the slide rule and the first calculating machines that used rotating cylinders to figure out the results of additions and multiplications appeared.

However, it was the computer that first carried out calculations that lay far beyond the abilities of the human brain. Machines were eventually capable of simulating deductive reasoning, a property of the mathematical mind. At that point some scientists felt they were crossing a frontier to which no machine of any kind had previously had access. Was this the right way to go? The development of information technology, with its exponential growth, was beginning to change paradigms that had been established for centuries. The first computational algorithms capable of proving theorems started to appear.

Critics of proofs done with a computer basically cite two reasons for calling this procedure into question. The first is that they are not possible to verify, as the program contains stages that can never be checked by any mathematician. The second is that the process is prone to errors, in both hardware and software. In most cases, these are random errors. One way of mitigating these effects is to use different programs in other machines to see whether they arrive at the same result.

But computers can only work with 1s and 0s. This implies some limitations, as, when numbers not expressed in base 2 are involved, they have to make approximations, which exposes them to potential errors. In 1991, David R. Stoutemyer carried out 18 experiments on calculations by computer programs that yielded incorrect results.

This has led many people to believe that new computational methods of research are best left for the experimental sciences, and not used for mathematics. However, nobody can say that mathematics can be performed using only a single method. 'Traditional' mathematical reasoning has never been free of errors throughout its history. On several occasions, wrong results have been considered correct for many years. Moreover, in our own time mathematics has reached such a high level of diversity and complexity that verifying a theorem can take years



It is calculated that the Cray supercomputer makes just one mistake every 1,000 hours of operation.

or, at best, can only be fully understood by just a few specialists. Lastly, there are many experts who now think that using computers as a research tool and for verifying theorems has given rise to a different way of viewing mathematics. It would not be unreasonable to suppose that one day Riemann's conjecture

MAXIMUM SECURITY

The United States government allows only certain cryptographic codes to be used on its territory and in Canada. There is also a ban on selling them outside these countries, unless it is to a financial institution. The unauthorised export of encryption standards is considered equivalent to weapons trafficking. Firms dedicated to the production of encryption programs store their secret codes in tablets equipped with sophisticated security devices. When they are broken open, their contents solidify into a shapeless mass on coming into contact with oxygen; if an attempt is made to view their content with X-rays or similar scans, whatever is written in them is converted to zeros.

may be proven by a computer. In any event, nobody can question the validity of computational methods used to find prime numbers and to check whether a number is prime. When we enter the world of computational algebra, terms like 'polynomial', 'deterministic polynomial' and 'probabilistic' are bandied about with great confidence, but they leave the uninitiated completely in the dark. Although not directly relevant here, it is useful to have some understanding of the notions encompassed by these terms.

When we speak of polynomial (or polynomic) time we mean the time taken by the machine to solve a particular algorithm. Suppose that we have an input variable we can call n. In general, when the algorithm uses a polynomial type of expression, such as $n^3 + 2n + 1$, we call this is a polynomial time algorithm (P), whereas, if exponential expressions were involved, we would be talking about a non-polynomial algorithm (NP). The basic idea, in general terms, is that polynomial algorithms have an acceptable operating time whereas exponential ones do not.

P versus NP

Computing may involve a series of problems that can be solved in a deterministic fashion. In other words the solution containing guarantees of its validity. For this we use polynomial algorithms that operate in polynomial time. The simplest example would be to perform additions or multiplications or to solve a large number of equations. In most cases, by using suitable algorithms, solution times can be kept within reasonable limits. All problems that can be handled in this way are known as P problems.

On the other hand, NP problems are those for which a non-deterministic solution can be found, by trying out solutions that may be correct. The time taken to solve this type of problem is very much shorter than that taken for P problems. It is clear that any problem that admits a deterministic solution in polynomial time is also a problem to which a solution can be applied for quick checking. In other words, any P type problem is also an NP type problem. However, having reached this point, we need to clarify the notion of an algorithm.

An algorithm can be compared to the method in a cookery recipe. It consists of a series of instructions that should be perfectly clear.

For example, to solve an equation such as $x - 2 = 8$, the solution algorithm would say something like:

1. Express the equation in terms of x (by moving all left-hand terms other than x over to the right-hand side): $x = 8 + 2$.

2. Add the terms on the right-hand side: $8 + 2 = 10$.

3. Write the solution: $x = 10$.

This is a P type problem, which would take a polynomial time to solve – it is a very simple one and would be quick to solve.

Of course, we could just try solutions like $x = 3$, $x = -2$, etc., and the computation time would be much shorter, as the only thing the program has to do is to enter a value in place of x and check if the solution is correct. It is non-deterministic because there is always a likelihood that the solution will be wrong. (We assume that we have some criteria to select the range of possible solutions, such as knowing that they should all fall between 9 and 11.)

The inverse of this question is as follows. If we have a test algorithm, can we guarantee that there is a polynomial algorithm that will allow us to solve the problem deterministically? This is almost the same as asking if we can be sure there is some kind of algorithm for seeking a solution in polynomial time.

This is the problem posed independently by Stephen Cook and Leonid Levin in 1971. If every P problem is NP, are there NP problems that are not P? This is considered the greatest challenge faced by modern computer science and is one of the Millennium Prize Problems established by the Clay Mathematics Institute. Anyone who solves it will win a million-dollar prize.

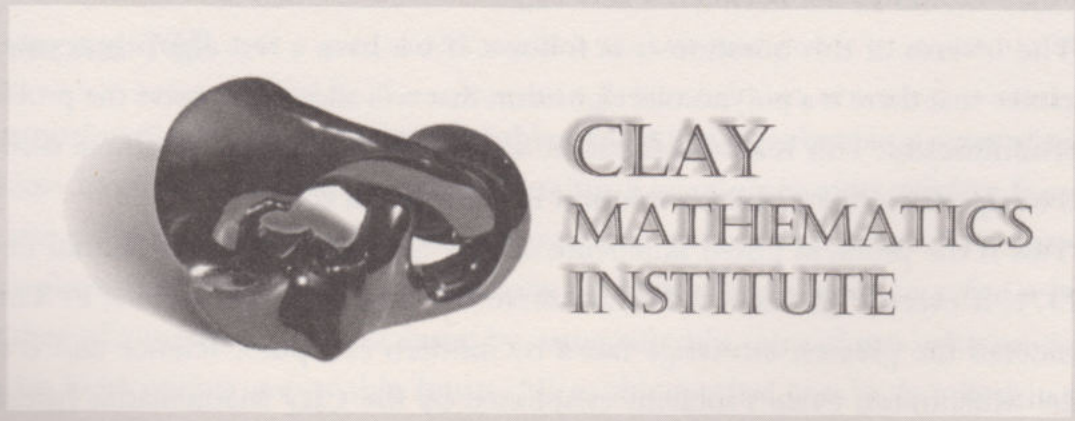
THE SEVEN MILLENNIUM PRIZE PROBLEMS

The Clay Mathematics Institute is a private, not-for-profit foundation established by Landon T. Clay, a multimillionaire entrepreneur from Boston. Its aim is to develop and disseminate mathematical knowledge.

On 25 May 2000, the institute announced the 'Millennium Prize Problems', with a total of seven million dollars to be paid out when seven problems, considered the most intractable in 20th-century mathematics, are solved. The problems can be solved one by one, that is, each will attract a prize of a million dollars (more than the Nobel Prize money) when it is solved. There are no time limits or age restrictions on candidates and no university background is required. The seven problems selected are:

1. The P versus NP problem.
2. The Riemann conjecture.
3. The Yang–Mills theory.
4. The Navier–Stokes equations.
5. The Birch and Swinnerton-Dyer conjecture.
6. The Hodge conjecture.
7. The Poincaré conjecture.

In view of the difficulty and significance of the problems proposed, Mr Clay's financial advisors doubted very much whether the Institute would ever have to pay out. However, in 2006, the Russian Grigori Perelman surprised everybody by solving the seventh and last problem, the Poincaré conjecture. Nevertheless, for personal reasons he rejected the Fields Medal awarded to him at the 25th International Congress of Mathematicians held in Madrid and he also turned down his Millennium Prize winnings.



Generating prime numbers

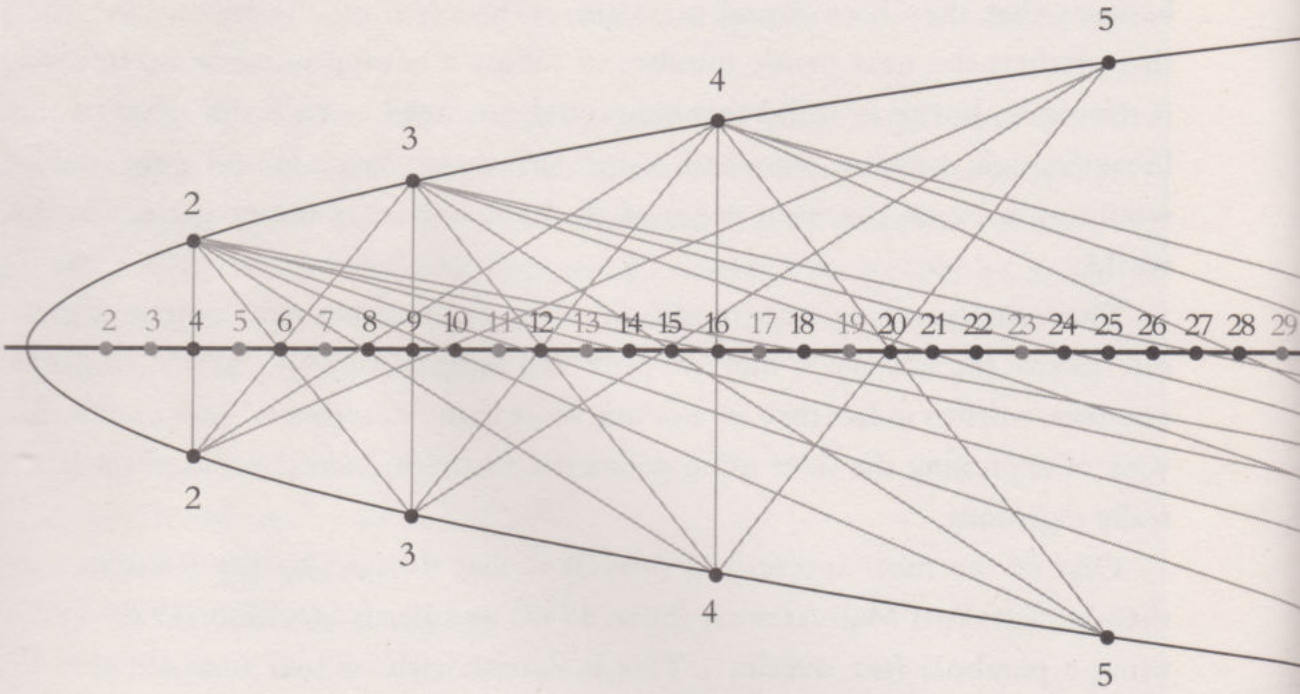
It is often the case that someone without much of a mathematical background believes that they have found a system or formula, usually on the Internet, that predicts the next prime number to follow a natural number n. However, it should be borne in mind that we would not need to seek out news of this breakthrough. Anyone who had found the magic formula and gone public would soon be on the front pages of all the newspapers and magazines in the world.

There are many geometric models for finding prime numbers. Sometimes these can deceive the unwary, as they are presented as formulae that can find all prime numbers, whereas in fact they are nothing more than Eratosthenes' sieve or different ways of expressing the sieve using geometric methods. Indeed, some of them are really ingenious.

One of the most interesting models is that devised by the Russian mathematicians Yuri Matiyasevich (born 1947) and Boris Stechkin (1920–1995) using a parabola (see overleaf). This is drawn with its two arms divided by a horizontal axis, with the sequence of natural numbers written along it. A perpendicular corresponding to the square of each number is then added, i.e. at the point +4 a perpendicular is drawn with the number 2 marked where it intersects both arms of the parabola. The geometric significance of the perpendicular is that it represents the product of $2 \cdot 2$; similarly, we draw another perpendicular from point 9 on the axis to symbolise the product of $3 \cdot 3$; and so on along the axis.

When all the numbers on the axis are represented by points on the parabola, each point on one arm is joined to all points on the other arm, that is, the point 2 on the upper arm is joined to 2, 3, 4, 5, etc., on the lower one. Each of these chords cuts the axis at the corresponding product of the two connected numbers – for example, the chord connecting 2 and 3 intersects at 6. If all possible intersections were made shown, the only points of the parabola free of intersections are the prime numbers. This is rather a satisfying example of a geometric sieve.

Sieves of an algebraic variety are more useful for obtaining quick computational algorithms. One of them is Atkin's sieve, designed by A.O.L. Atkin and Daniel J. Bernstein, which enables all prime numbers less than or equal to a given natural number to be found. In some senses it is an improved version of Eratosthenes' sieve.

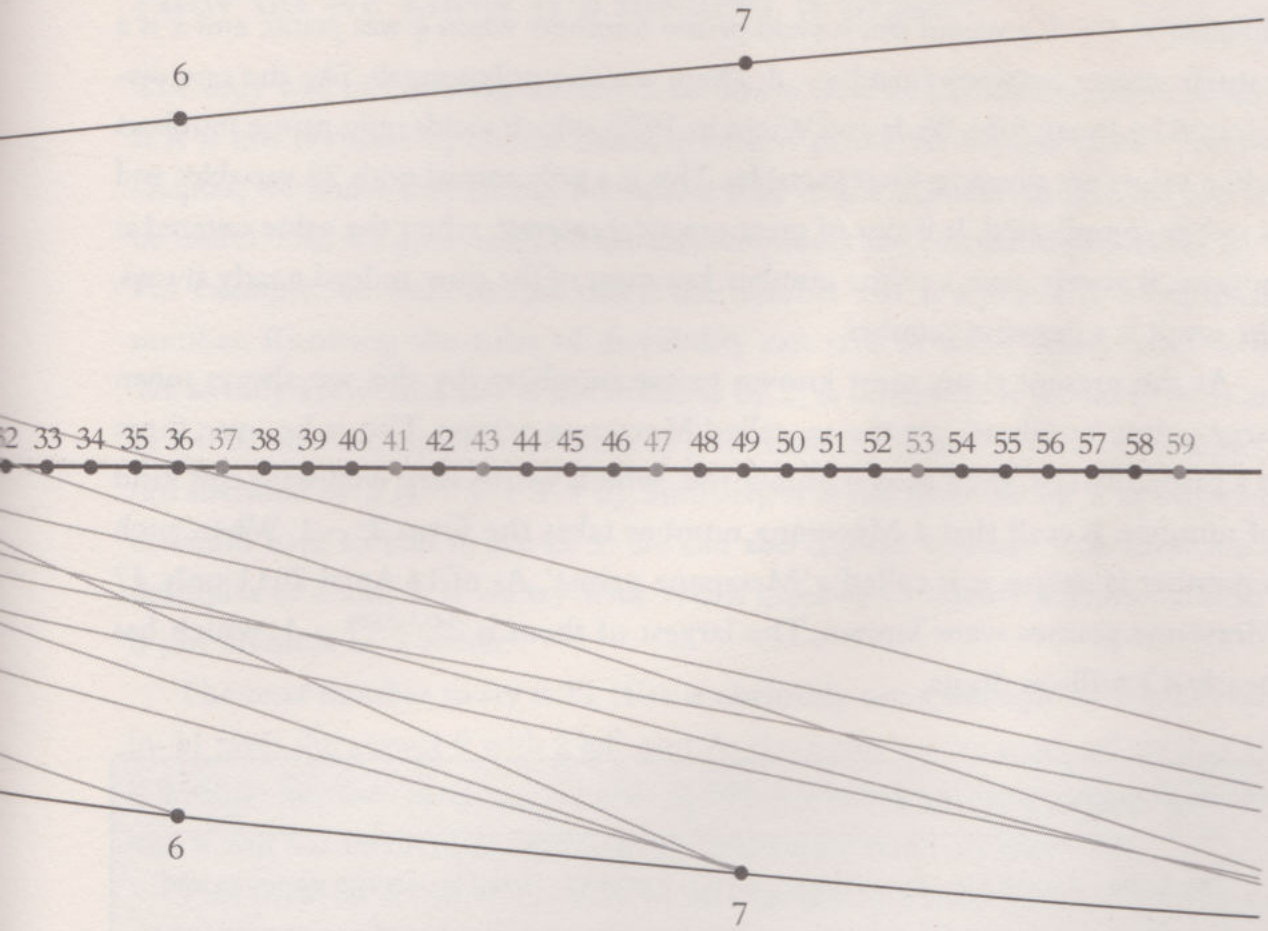


The geometric sieve devised by Yuri Matiyasevich and Boris Stechkin for finding prime numbers, which appear in grey in the illustration. Note that no chord passes through these numbers.

When we say 'improved' we really mean 'updated', as Atkin's sieve, arithmetically speaking, contains some shortcomings compared with Eratosthenes', given that it requires some preparation and does not eliminate multiples of prime numbers but only multiples of the squares of primes.

We now know that the ideal thing would be to find a formula that associates each natural number n with the nth prime number. We have seen how mathematicians have sought this formula for at least 3,000 years. What we do have are functions that allow us to calculate prime numbers in a practical way.

For example, it is proven (Wilson's theorem) that p is a prime number if and only if $(p-1)! \equiv -1 (\text{mod. } p)$. However, as we touched on previously, any formula containing factorials is infeasible when we use an algorithm in a computer, because rapid growth of the function would make computing times too long.



There are also polynomials that 'generate' prime numbers, like that used by Euler to draw up a list of 40 primes using the function $f(x) = x^2 + x + 41$, which generates prime numbers when values of x are entered. For example:

$$x = 0 \quad f(0) = 0 + 0 + 41 = 41$$
$$x = 1 \quad f(1) = 1 + 1 + 41 = 43$$
$$x = 2 \quad f(2) = 4 + 2 + 41 = 47.$$

However, the formula begins to fail for values of x greater than 41. For example, $x = 41$ results in composite number:

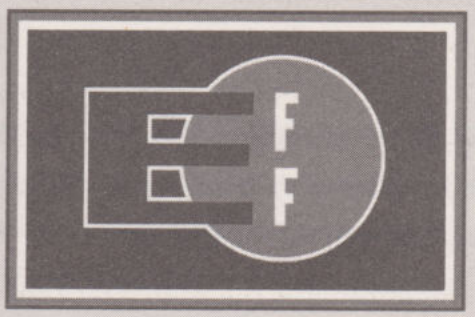
$$x = 41 \quad f(41) = 1{,}681 + 41 + 41 = 1{,}763.$$

Euler proceeded to study this polynomial and came to the conclusion that the formula $x^2 - x + q$ would often yield prime numbers when q was prime and x is a natural integer between 0 and $q - 2$. There are also polynomials, like the one discovered by Jones, Sato, Wada and Wiens in 1976, which yields only prime numbers when values are given to their variables. This is a polynomial with 28 variables and is rather complicated. It is not of great practical interest: when the value entered is positive, it always gives a prime number, but most of the time, indeed nearly always, the result is a negative number.

At the present time, most known prime numbers (by this we always mean large prime numbers) are the so-called Mersenne primes. This is because there is a primality test, the Lucas–Lehmer test, which works very well with this kind of number. Recall that a Mersenne number takes the form $2^n - 1$. When such a number is prime, it is called a 'Mersenne prime'. As of 14 April 2011 only 47 Mersenne primes were known. The largest of them is $2^{43,112,609} - 1$, which has nearly 13 million digits.

THE GIMPS PROJECT

The Great Internet Mersenne Prime Search is a project started by George Woltman and consists of a collaborative network in which the personal computers of people participating in the project (anyone can sign up) operate in parallel and achieve capacities much greater than those offered by any current supercomputer. Each participant in the scheme installs the appropriate software supplied by the organisation, and their computer works on the problem during periods of down time. The project started to run in 1997, and by August 2009 a total of 12 new Mersenne primes had been discovered. The Electronic Frontier Foundation (EFF) offered a prize of $150,000 for the first person who discovered a Mersenne prime with at least ten million digits. On 23rd August 2008 the prize was awarded to Edson Smith of the Mathematics Department. at University of California, Los Angeles for discovering the number $2^{43,112,609} - 1$. On 7th January, 2016, GIMPS celebrated its 20th anniversary with the discovery of the largest prime number $274,207,281 - 1$.



The logo of the Electronic Frontier Foundation.

How do we know if a number is prime?

The only way of knowing for sure is to divide it by all the numbers preceding it. If it is not divisible by any of them, then it is prime. As we saw in the previous chapter, we know that taking the square root of the number concerned can also be useful. This is a good method for small numbers and calculations done by hand. For example, we want to find out if the number 101 is a prime or a composite number. Knowing the rules of divisibility can save us some unnecessary work. We already know that 101 is not divisible by 2, as otherwise it would have to end in a 0 or an even number. Neither is it divisible by 3, as the sum of its digits is not divisible by 3 (1 + 0 + 1 = 2). Similarly, it is not divisible by 5, as otherwise it would have to end in a 0 or 5. We can also discard 4, 6 and 9, as these are all multiples of 2 and 3. If we try with 7, this gives us 14 with 3 left over, and so it is not divisible by 7 either.

The next number to try is 11 (101 is obviously not a multiple of 10). Division by 11 gives the answer 9 with 2 left over. At this point we can stop and say that 101 is a prime number, as the square root of 101 is approximately 10, which assures us that it will not be divisible by any of the remaining numbers up to 101.

This method is called 'division by trial' and is the simplest and most secure of all. The problem is that it is not feasible for very large numbers, not even if a computer is used. Note that a 50 digit number would require divisions using numbers up to 25 digits in length, which would correspond more or less to its square root.

A computer able to perform a billion divisions per second would take somewhat longer than 300 million years to complete the task, and by then it is likely that humanity will have disappeared from the face of the Earth! Anyway, the point is that the method works well enough for a composite number if one of its factors is not too large. We should bear in mind that, given any number n, the probability that the number 2 is a factor is 50%, that 3 is a factor 33% and so on.

On the other hand, modern computers have been making significant advances in speed and memory capacity so that seeking a prime number in a long list is sometimes more efficient than the complicated process of determining whether a given number is prime. It's important to take steps to ensure that the algorithms underlying computer software are sound and reliable, or they may fail to give reliable results. In big calculations, a small error at one stage can propagate through and lead to huge problems later on.

Pseudoprimes

Fermat's little theorem is one of the most frequently used in primality tests. This theorem states: "If p is prime there is no base a with $a < p$ (a and p being primes themselves), so that $a^{p-1} - 1$ gives a remainder other than zero when it is divided by p".

The theorem has its limitations because, as we have seen, it postulates a necessary but insufficient condition. For example, if we take $p = 7$, we are saying that $3^6 - 1$ is divisible by 7. There is no guarantee that 7 is a prime number (we know it is prime because it is a small number taken to simplify the example, but we should imagine we are dealing with large numbers). However, if we take $p = 8$, we are saying that $3^7 - 1 = 2{,}186$ which is not divisible by 8, assuring us that 8 is not prime (without having to find any of its factors).

We know that any number that does not pass the test for a given base is a composite.

If, on the other hand, the number passes the test and is prime, we describe the number of the base as 'false'. We can then go on testing. The probability of finding false numbers is reduced by ½ with each test, so that the probability that the number is prime keeps increasing.

A number p that, although not being prime, passes a test for a base a is said to be a pseudoprime for that base. A more general definition of a pseudoprime is: 'A number is said to be a pseudoprime if it passes a prime-number test and turns out to be composite'.

The matter gets more complicated for numbers that pass tests for any base a and are not prime. For example, the number 561 passes the test for any base and yet is a composite number ($561 = 3 \cdot 11 \cdot 17$). These numbers, discovered by the US mathematician Robert Daniel Carmichael (1879–1967) are known as 'Carmichael numbers'. To date only 2,163 Carmichael numbers are known, and they are found among the first 25 billion natural numbers. They all have at least three prime factors.

There are 16 Carmichael numbers smaller than 100,000, namely:

561; 1,105; 1,729; 2,465; 2,821; 6,601; 8,911;
10,585; 15,841; 29,341; 41,041; 46,657; 52,633;
62,745; 63,973 and 75,361.

Carmichael numbers are also called 'absolute pseudoprimes'.

Certification methods

Today there are two types of algorithm used to determine whether a number is prime: the deterministic polynomial and the probabilistic polynomial.

The first of these establishes for certain whether a number is prime, but it takes a long time to perform. The second is quicker, but there is some uncertainty in the result.

The most widely used method is called the 'Miller–Rabin method', a version of Fermat's primality test that is based on the Riemann conjecture. It is a probabilistic polynomial, but the probability of it producing an error is between $1/10^{50}$ and $1/10^{80}$ and so, in practice, it can be assumed to be wholly accurate.

On 6 August 2002, three researchers from the Technology Institute in Kanpur (India), M. Agrawal, N. Kayal and N. Saxena, published a deterministic method in polynomial execution time based on a generalisation of Fermat's little theorem. It is known as the AKS primality test:

$$n \text{ is prime } \Leftrightarrow (x-a)^n = x^n - a \text{ and the range } \frac{\mathbb{Z}_n[n]}{x^r - 1} .$$

AKS was the first primality-proving algorithm to be simultaneously general, polynomial, deterministic and unconditional. Previous algorithms had achieved three of these properties, but not all four.

Despite this, the method most used continues to be the probabilistic polynomial in view of its shorter run time.

Most web browsers include an encryption algorithm capable of using this kind of method to find large prime numbers of up to 2,048 bits.

Today, the three cryptographic security systems used are RSA, DSA (Digital Signature Algorithm) and ECDSA (Elliptical Curve Digital Signature Algorithm).

No expert doubts the security provided by each of these three systems. The difference between them resides in the codes they use. The security provided by the 2,048 bit codes in the first two is equivalent to using 224 bits in the third, whereby the calculation time is considerably reduced. Whereas sub-exponential algorithms are used in the first two, the third uses an exponential type and has yielded the best results so far.

NUMERICAL CURIOSITIES

The number 313 appears on the numberplates of cars driven by Donald Duck. It has the curious property of being palindromic – it can be read from left to right and right to left – in both base 10 and base 2, and it is the only three-digit prime number with this property: 313 (base 10) = 100111001 (base 2). Furthermore, 100,111,001 in base ten is prime.

There are many prime numbers with strange properties. For example, some are known as 'repunit' (a word coined from the term 'repeated unit'), which consist of long sequences of 1s. The number 11111111111111111111 (23 ones) is a prime. In principle, they are just that – curiosities – although one day these numbers may become part of a theorem or conjecture with some mathematical value. Another curious sequence of this type is the one based on the number 91, which is a composite ($91 = 13 \cdot 7$), but when sequences of 9s and 0s are inserted into the middle of the number, the resulting number alternates between prime and composite:

9901 prime
999001 composite
99990001 prime
9999900001 composite
999999000001 prime
99999990000001 composite
9999999900000001 prime
999999999000000001 composite

Unfortunately, the next one, 99999999990000000001, is also composite!

The story continues...

We have seen how mathematicians like Mersenne, Fermat and, on occasion, even Euler himself, were looking for practical number tools. This often undermined the consolidation of pure theory to some extent. Proofs were barely alluded to, but the results went on being used. Gauss turned the page in the history of mathematics by insisting that producing rigorous proofs should be the overriding goal. However, with prime numbers, we seem to have re-adopted the empirical approach. We use unproven theorems and approve results trusting that the chance of them containing

an error is very small. We act like Fermat, but without even the need of concealing a hypothetical proof. We have reached this point because of, on the one hand, our enormous capabilities with computer algorithms, and on the other because of our great need for large prime numbers.

On a purely theoretical level it could be said that prime numbers continue to resist the efforts of mathematicians. Their history is, to a great extent, a history of failure. The greatest success we have had is with Riemann's zeta function, but we are painfully aware that this is only a partial success. Euler, who was a great mathematical visionary, was not particularly optimistic about the chances of eventually understanding these elusive numbers:

> Mathematicians have long tried in vain to discover some pattern in the sequence of prime numbers, but I have reason to believe that this is a mystery that the human mind will never be able to penetrate.

Appendix

Proofs

Proof of the fundamental theory of arithmetic

The theorem states that every natural number other than 1 can be expressed as a single product of prime factors.

First we should explain why 1 is not considered a prime number. There are several reasons but the most obvious is that, if this were not so, the theorem would not hold, as the number 1, being prime, could be factorised in a number of ways:

$$1 = 1 \cdot 1 = 1 \cdot 1 \cdot 1 = 1 \cdot 1 \cdot 1 \cdot 1 = \ldots$$

With this proviso then, we can prove the theorem in two stages. The first demonstrates that the number can be reduced, and the second that there is only one way of doing so.

The first part involves *reductio ad absurdum*. Suppose that n is the smallest number that cannot be broken into prime factors. We know that this cannot be 1 because we dismissed that possibility when stating the theorem. Neither can it be a prime number because this can only be reduced to itself. Therefore, it has to be a composite number of the form $n = a \cdot b$, where a and b are smaller than n. But as n is the smallest number satisfying the condition that the number should not be reducible to prime factors, this means that a and b are indeed reducible and that n must also be so, thereby arriving at a contradiction.

The second part of the proof is based on the following result,

If p is a prime number that divides a product of factors, it must necessarily divide into one of these factors. This result can be proven by using Bézout's identity. Suppose that a natural number greater than 1 can be reduced to prime factors in two ways: we take a prime number p from the first reduction. This number must necessarily divide into the second reduction and, therefore, into one of its factors. We select the factor into which it divides and, as this is a prime factor, it has to be equal to p. We have now found two factors with the same reduction. We now take another prime number and continue the process until we see that the prime factors appearing in both reductions are all the same.

Proof of Fermat's little theorem

Expressed as an identity, as in chapter 5, the theorem states that: "If p is a prime number, then for every natural number a, $a^p \equiv a \pmod{p}$."

The theorem is equivalent to proving that p divides into $a^p - a$.

Let's prove the theorem by using the inductive method for a. In other words, we suppose that it is true that for a natural number a, and we then show that it is also true for $a + 1$.

Therefore, we start with the hypothesis that p divides into $a^p - a$. According to Newton's binomial expansion:

$$(a+1)^p = a^p + \binom{p}{1}a^{p-1} + \binom{p}{2}a^{p-2} + \ldots + \binom{p}{p-1}a + 1.$$

Moving a^p and 1 over to the left-hand side, we have

$$(a+1)^p - a^p - 1 = \binom{p}{1}a^{p-1} + \binom{p}{2}a^{p-2} + \ldots + \binom{p}{p-1}a.$$

The factor p is in all the terms on the right-hand side, and so p divides into the right-hand side of the equation and, hence, into the left-hand side, $(a + 1)^p - a^p - 1$, too.

Because according to the inductive method, p divides into $a^p - a$, we can say that it also divides into the sum

$$\left[(a+1)^p - a^p - 1\right] + a^p - a.$$

This sum, suitably rearranged, can be expressed in the form

$$\left[(a+1)^p - a^p - 1\right] + a^p - a = (a+1)^p - (a+1).$$

Hence, we can state that it is also true for $a + 1$ and, therefore, the theorem is proven.

Bibliography

BENTLEY, P. J., *The Book of Numbers*, Ontario, Firefly Books, 2008.

HARDY, G. H., *A Mathematician's Apology*, Cambridge, Cambridge University Press, 2012.

HARDY, G. H., *Collected Papers of Srinivasa Ramanujan*, Cambridge, Cambridge University Press, 2015.

IFRAH, G., *The Universal History of Numbers*, London, The Harvill Press, 1998.

KANIGEL, R., *The Man who knew Infinity: A Life of the Genius Ramanujan*, New York, Washington Square Press, 1991.

KLINE, M., *Mathematical Thought from Ancient to Modern Times*, (3 Volumes), USA, Oxford University Press, 1990.

PICKOVER, C. A., *Wonders of Numbers*, USA, Oxford University Press, 2002.

SAUTOY, M. DU, *The Music of the Primes: Why an Unsolved Problem in Mathematics Matters*, London, Harper Perennial, 2004.

STEWART, I., *From Here to Infinity*, Oxford University Press, 1996.

SZPIRO, G., *Poincare's Prize: The Hundred-Year Quest to Solve One of Math's Greatest Puzzles*, London, Plume, 2010.

Index